DOCUMENT RESUME

ED 284 165 CS 008 761

AUTHOR Simons, Herbert D.; Ammon, Paul R.
TITLE The Language of Beginning Reading Texts. Final Report.
SPONS AGENCY National Council of Teachers of English, Urbana, Ill. Research Foundation.
PUB DATE Apr 87
NOTE 70p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Basal Reading; *Beginning Reading; Educational Assessment; Evaluation Methods; Instructional Effectiveness; Language Usage; Learning Strategies; *Miscue Analysis; Primary Education; *Readability; Reading Ability; Reading Achievement; *Reading Comprehension; *Reading Processes; Reading Research; Skill Analysis; Student Evaluation; Teaching Methods; *Textbook Content; Theory Practice Relationship

ABSTRACT

To weigh the negative effects of "primerese" against its supposedly improved readability, a study compared the reading of controlled basal stories with that of the same stories minimally rewritten into more natural language. Participants were 64 white, lower middle class first grade students selected from five classrooms (all with basal instruction programs) in two communities. The children were paired by word recognition scores, each member reading alternate versions of the same four stories over a two-week period. They were tested for comprehension and for oral reading miscues. In comprehension, subjects with high word recognition scores profited from changes in the revised versions, while low word recognition subjects were not affected. Students reading the rewritten version answered a majority of questions correctly more often, suggesting that reduced dependence on pictures and natural language conventions facilitated inference, coordination of different parts of the text, and a greater use of expectations based on natural language in top-down processing of the texts. Analysis of oral reading miscues indicated that while the rewritten versions had no significant effect on rate of reading or word miscues, they reduced the proportion of graphically based miscues for high word recognition subjects, who consistently produced fewer punctuation miscues, suggesting that the rewritten texts were more in keeping with expectations children derive from naturally occurring language patterns. Tables, references, texts and test questions are included. (JG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED284165

THE LANGUAGE OF BEGINNING READING TEXTS

Principal Investigators

Herbert D. Simons
Paul R. Ammon

Research Associate
Charles Elster

Research Assistants
Linda Edelman
Judy Jones
Sharon Trattner

University of California, Berkeley.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"See Spot. See Spot run. See Spot run and play...' What IS this garbage?"

Final Report: National Council of Teachers of English Research Foundation
April 1987

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Herbert D. Simons

Paul Ammon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

CS008761

Acknowledgments

We wish to thank all those people who helped make this project possible. The most important people on the project were our research associate Charles Elster, who did a superb job of overseeing and participating in all phases of the project, and our research assistants Linda Edelman, Judy Jones and Sharon Trattner whose diligence and willingness to work long hours in collecting, transcribing, and coding the data went well beyond the call of duty. We owe an intellectual debt to Roger Shuy whose critique of controlled text language stimulated our interest in the project. The work of Kenneth and Yetta Goodman also played major role in our use of reading miscues to interpret the children's reading performance. Special thanks go to Yetta Goodman for taking a personal interest in the project and providing support and encouragement. We also want to thank Robin Lakoff for helping to analyze the primerese features in the texts used and Wallace Chafe for providing advice on some on the repunctuation analysis. Special thanks go to the teachers Mrs. Amess, Mrs. Huffman, Mrs. Maddox, Ms. Turner and Mrs. Waters in whose classrooms the research was conducted. We also want to thank the principals of their schools Mr. Mark Stott, Mr. Frank Murray and Dr. Sam Tracas for their cooperation. Finally, we wish to thank the National Council of Teachers of English Research Foundation for funding the project and the Committee on Research of Berkeley campus of the University of California and Dean Bernard Gifford for providing the supplemental funds needed to complete the research.

Herbert D. Simons

Paul R. Ammon

TABLE OF CONTENTS

Section		Page
I.	The Language of Beginning Reading Texts	1
II.	Method	7
	Design	7
	Subjects	7
	Procedures	8
	Stories	9
	Word Recognition Test	16
	Comprehension Test	16
	Reading Miscue Coding	16
	Approach to Quantitative Analyses	17
III.	Results	18
	Comprehension Questions	18
	Quantitative Analyses of Oral Reading	20
	Qualitative Analysis of Reading Miscues	38
IV.	Natural Language Expectations	40
	Editing Miscues	40
	Repunctuation Miscues	45
V.	Primerese Expectations	48
	Repeated Structure Miscues	49
	Miscues Involving Dialogue Carriers	51
	Educational Recommendations	53
	Conclusions	55
VI.	References	57
	Appendix A (Texts and Comprehension Questions)	60
	Appendix B (Word Recognition Text Word List)	64

LIST OF TABLES

Table 1. 15
Number of Words, Different Words, and Sentences
in Original (O) and Rewritten (R) Versions of Stories

Table 2. 19
Mean Comprehension Scores for Original and Rewritten
Version Within Low and High Word Recognition Groups

Table 3. 22
Mean Reading Rates and Miscue Rates for Students
With Low (L) and High (H) Word Recognition Scores

Table 4. 23
Mean Reading Rates and Miscue Rates for
Original (O) and Rewritten (R) Versions

Table 5. 24
Mean Miscue Rates in Significant Interactions
of Text Version and Word Recognition Level

Table 6. 26
Mean Proportions of Substitution, Omission, Insertion, and
Nonsense Miscues for Students with Low (L) and High (H)
Word Recognition Scores

Table 7 27
Mean Proportions of Substitution, Omission, Insertion, and
Nonsense Miscues for Original (O) and Rewritten (R) Versions

Table 8. 28
Mean Proportions of Syntactic, Semantic, and Graphic Miscues
for Students with Low (L) and High (H) Word Recognition Scores

Table 9. 30
Mean Proportions of Syntactic, Semantic, and Graphic Miscues
for Original (O) and Rewritten (R) Versions

Table 10. 31
Mean Proportion of Graphic Miscues in Significant Interactions
of Text Version and Word Recognition Level

Table 11. 33
Mean Miscue Rates on "New" Words and All Words in the
Rewritten Version for Students with Low and High Word Recognition Scores

Table 12. 34
Mean Proportions of Syntactic, Semantic, and Graphic Miscues
on "New" Words and All Words in the Rewritten Version for
Students with Low Word Recognition Scores

Table 13. 36
Mean Rates of Phonic and Punctuation Miscues for
Students with Low (L) and High (H) Word Recognition Scores

Table 14. 37
Mean Rates of Phonic and Punctuation Miscues for
Original (O) and Rewritten (R) Versions

The Language of Beginning Reading Texts

The texts used to teach beginning reading often employ a unique form of discourse. This form of discourse, commonly known as "primerese," is like neither spoken language nor well formed written language. It is the result of attempts to simplify reading texts in order to make learning to read easier. This simplification involves vocabulary and sentence length controls, and the extensive use of dialogue and illustrations. The vocabulary controls involve limiting the stories to a small set of high frequency words. The sentence length controls involve limiting the number of words per sentence.

Primerese is usually justified by citing the problems that beginning readers have when they encounter texts with a large number of words that they cannot decode. Such texts make fluent reading difficult and interfere with comprehension. Thus, in effect, the texts are not "meaningful." The concern for heavy vocabulary burden stems from the readers used in the first part of this century-- readers in which there was very little control over the number of different words introduced. The use of these uncontrolled texts meant that children in the primary grades encountered many words that they did not know. According to the proponents of primerese, this problem has been overcome by reducing the vocabulary burden, which involved reducing the overall number of different words, limiting their introduction to one or two new words a story, and repeating the words often. Meaningfulness is increased by using words that are in children's speaking vocabularies, using dialogue because it is closer to their oral language, employing illustrations rather than words to carry the meaning, and describing familiar experiences. Frequent repetition is employed to provide the practice necessary to learn to recognize the words as well as sound-letter correspondences. Short sentences in conformance with readability formulas are also used to decrease the processing burden. According to readability research, short sentences make a text easier than longer sentences. All of these controls reduce the vocabulary burden and allow readers to concentrate on meaning.

These primerese controls result, however, in texts that are characterized by their redundancy and awkwardness, due to frequent repetition of words in places where they are odd or unnatural, short sentences in which explicit connectives between sentences have been left out, use of exophoric reference that makes the interpretation of the text dependent upon the illustrations, and malformed stories in which basic elements have been omitted and/or are carried by the illustrations.

The stilted and unnatural language of primerese, particularly in first grade readers, has been criticized by almost everyone including even some of those reading educators associated with the publication of the readers themselves. The words "See Dick run. See Jane run" are sometimes quoted with nostalgia, but more often they are cited as the archetypical example of the awkward nature of primerese. It has been pointed out that the controls imposed on reading texts, rather than simplifying them, may make them more complex (Beck, McKeown & McCaslin, 1981; Gourley, 1978; Shuy, 1981; Shuy & Larkin, 1978; Goodman, 1986). By focusing only on vocabulary and sentence length, the authors of primerese texts ignore the syntactic, semantic, pragmatic, and discourse levels of language. These other levels of language

may actually become more complex and less coherent as a result of the superficial primerese controls. The result is that the language of these texts becomes less "meaningful" and is inconsistent with children's expectations based on their knowledge of the world, and of lexical, sentence, discourse, and story structure (Goodman, 1986; Gourley, 1978; Smith, 1982; Shuy, 1981). It has been claimed that this inconsistency makes it harder for beginning readers to read primerese texts than more natural language texts, and therefore also to learn how to read. This is the working hypothesis for this study.

This claim about the disadvantages of primerese finds general support in the work of cognitive psychologists and linguists who have shown the importance of background knowledge and expectations in learning and comprehending (Anderson, Spiro and Montague, 1977; Schank and Abelson, 1977). Empirical support for the criticisms of primerese can also be found in studies in which primerese texts have been compared to less controlled texts. If primerese disrupts children's expectations about texts, then texts written with primerese vocabulary controls relaxed should be more consistent with children's expectations and therefore easier to read. Bridge, Winograd, and Haley (1983) conducted an experiment to test this hypothesis and found that first grade children learned to recognize more words as a result of reading predictable texts without vocabulary control than from reading primerese texts. Brennan, Bridge, and Winograd (1986) also found that second grade students had better recall of a rewritten, well formed story than the original, malformed basal story. Beck, McKeown, Omanson, and Pople (1984) found that basal stories which were rewritten to make them more coherent produced better comprehension in third grade students than the original basal stories. And Beck, Omanson, and McKeown (1982) found that third grade students had better comprehension of rewritten stories in which relevant background knowledge was added and important information was highlighted. Feldman (1985) found that first grade children's recall of primerese stories rewritten to conform to story grammars was superior to the recall of the original story. It is important to note, however, that in all of these studies the content, as well as the language and structure, was different between the versions, thus making it difficult to know precisely what was responsible for the superiority of the less controlled texts.

Other research supporting the criticisms of primerese has examined specific aspects of the justification for primerese. Evidence on the necessity of controlling vocabulary to limit exposure to too many unfamiliar words come from two studies by Gates (1961, 1962). He tested students in grades 2 and 3 on words already covered in their readers (old words) and words that would appear in subsequent years in their readers but had not appeared in the readers in their current grade (new words). He found that students knew a substantial number of new words, often as many as the old words. This finding raises questions about the rationale of limiting the number of new words. It appears from the Gates research that children may already be able to read many words that do not appear in controlled texts, so that these words would not really add to the vocabulary burden.

Finally, Gourley (1984) examined the effects on readers' miscues produced by several specific linguistic features of primerese, including repeated noun phrases, lack of connectives, improper use of definite and indefinite articles and verb tense, unmotivated patterned repetition, and the use of illustrations. She found that these features produced miscues.

General support for primerese comes from traditional learning theory, which stresses the important function of repetition and practice in learning, and from readability research, which suggests that short sentences and easy words are the most important features of easy texts (Chall, 1958; Klare, 1963). Support also comes from the almost universal experience of primary school teachers who see students struggling to read stories with too many words they do not know.

More direct experimental evidence comes from the research of Gates (1931), who studied first grade classes in which new words were introduced in the readers at different rates. In one study he compared the word recognition and comprehension test scores of classes in which the introduction of new words was restricted to 1 new word in 150 running words (1-150) with classes using traditional basal readers that had a 1-16 new word introduction rate. He found that the controlled vocabulary (1-150) treatment produced better reading only for the low I.Q.(60-80) students. The heavy vocabulary burden (1-16) actually benefitted the high I.Q. (100-120) group while the medium (80-90) I.Q. group showed no differences. In a second experiment, he compared the same controlled (1-150) to less controlled (1-60) materials for low I.Q. students (mean = 73), and found the less controlled materials produced generally better word recognition. Thus the evidence here in favor of vocabulary control is not overwhelming, but it does suggest that some degree of vocabulary control helps to overcome the vocabulary burden problem for low I.Q. students. However, even for these students, too much vocabulary control can be counterproductive.

Gates and Russell (1938a) compared students matched on M.A. and readiness scores in classes which used only basal readers and in classes which used basals supplemented by a large amount of additional reading materials. The supplemental material increased the number of repetitions per word, which made the words more familiar and presumably reduced the vocabulary burden. They found no statistically significant differences between the basal and the supplemental materials group for the groups as whole and for subjects in high, medium, and low ability groups. In a further study of the same subjects, Gates and Russell (1938b) studied three matched groups of children who were introduced to three degrees of vocabulary control. They found that, on measures of word recognition and paragraph reading, the group that had been introduced to the smallest and medium numbers of different words scored higher than the students who had encountered the largest number of different words. When the subjects were divided into high, medium, and low reading readiness groups, Gates and Russell found that the low readiness group benefitted most from the limited set of words, and that the medium group had some benefit also, while for the high readiness group the most controlled vocabulary actually produced poorer reading. This study, as well as the others we have cited, suffered from problems in not controlling teacher differences and other classroom variables. In addition, the treatments included workbooks and other

reading materials, so they were not limited to the study of vocabulary control per se. However, even if these problems are put aside for the moment, the evidence only provides support for controls for low and perhaps average ability students. But this qualification does not appear to have influenced the use of controls in basals for all students. Basals appear to be aimed at the lowest common denominator, apparently on the assumption they will help the low ability students while not harming high ability students.

The other major primerese control, that of sentence length, appears to come from readability research which has shown that, on the average, shorter sentences are easier to read than longer ones. This research has had and continues to have a great deal of influence on the publishers of basal readers and other types of reading materials for children. The readability formulas upon which the primerese constraint of short sentences is based have been subject to extensive criticism with respect to their inadequacies in evaluating sentence and discourse level features. These criticisms raise questions about this constraint (Bruce, Rubin, & Starr, K. 1981; Davidson, 1984). However, even if the validity of readability formulas were acceptable, their use as a guide in rewriting was always considered problematical by the readability researchers themselves (Chall, 1958; Klare, 1963).

Another feature of primerese is the extensive use of illustrations. Beyond the motivation they presumably provide, they are often used to carry concepts and parts of the story that are not in the text (Elster & Simons, 1985). This makes the stories picture dependent, in that the illustrations are frequently necessary to understand the story. There has been some research on the usefulness of pictures in word learning (Samuels, 1970), generally suggesting that they do not help, but no studies have been conducted on the specific issue of the picture dependence feature of primerese.

The extensive use of dialogue in primerese, which appears to be motivated by a desire to mimic children's oral language, has not been the subject any research that we know of.

Overall, the arguments and evidence of the critics of primerese seem more persuasive than those of its defenders. However, the issue is far from settled. The direct experimental evidence has some problems, as mentioned above. The accumulated experience of first grade teachers suggests that there is a need for some limits on vocabulary, which has been the major motivation for primerese. The problem is that the vocabulary and other related controls have a number of negative effects on the language of the stories that may make them harder to read, even though the words are relatively easy. Thus the question is whether what is gained in reading ease by vocabulary and other controls compensates for negative sentence and text consequences.

The present study was designed to help answer this question by comparing first grade children's reading of controlled basal stories with their reading of these same stories rewritten into more natural written language with controls relaxed in relatively modest way. Our goal in rewriting the original versions was to eliminate unnatural features that were the result of conventional controls, and to replace them with features that were more

consistent with written book language outside the domain of basal readers. The content of the stories was the same in both versions, but some of the content was implied or carried in the illustrations in the original versions, while in the rewritten version it was also part of the text. The original language was left intact whenever possible. Thus our rewriting was minimal and did not attempt to change the stories into models of children's literature.

To assess the effects of these changes, we examined both comprehension and oral reading miscues. Comprehension was chosen as one major dependent variable because it is the goal of the reading process. A detailed analysis of miscues was included because of our interest in the reading process itself, and in the ways different versions might affect it. Unlike responses to comprehension questions, reading miscues provide fairly direct evidence as to the manner in which readers process a text as they read. Variations in the reading process in response to different text versions are of interest not only for their potential effects on comprehension of the immediate text, but also for their implications regarding the kinds of strategies children might learn to adopt more generally when reading.

Our treatment of reading miscues is based on the premise that readers form expectations about what they are about to process while reading, and that the fit (or lack thereof) between these expectations and the actual text can affect the accuracy or efficiency of the reading process, and perhaps the very nature of that process as well. We think of the reader's expectations as momentary states that are based jointly on the preceding context (as perceived by the reader) and on background knowledge about language and the nonlinguistic world. The linguistic knowledge giving rise to a reader's expectations can include semantic, syntactic, pragmatic, and discourse levels of analysis. Such knowledge tends to participate actively in the formation of expectations, unless its influence is otherwise inhibited or suppressed.

We assume, further, that all readers--including young novices and older experts--have such expectations. However, the extent to which these expectations actually affect word recognition is a matter of some controversy. Some reading theorists have argued that expert readers decode so quickly and accurately that highly automatized "bottom-up" processing of graphic cues must be the principal basis on which the words in a text are recognized (Stanovich, 1980). However, even expert readers occasionally make miscues which suggest that "top-down" reader expectations play a role in word recognition. Moreover, such processes are central to current "constructivist" views of reading comprehension, and it is difficult to believe that they would not also influence word recognition. That seems especially likely to be the case with beginning readers who have attained some degree of reading fluency but not the kind of automaticity that comes with further practice. In any event, the validity of our assumptions in this regard can be tested against our findings.

A child reader's expectations are related to the child's own spoken language and to the spoken and written language that the child has been exposed to, but the child's expectations are not directly predictable from any of these observable sources. That is because the child does not simply copy linguistic input, but rather transforms it in the process of assimilating it, and because both expectations and oral productions are constructed--on the spot--and are not simple reflections of what the child already knows. Consequently it would be difficult to predict exactly which texts would best fit a child's expectations, even if one knew a fair amount about the child's past experience and present ways of speaking. Nevertheless, it does not seem unreasonable to assume that the linguistic knowledge underlying children's expectations at least represents a good approximation to the intuitions adults have about well formed natural written language texts, so that our attempts to make primerese texts more natural for children can be guided by such adult intuitions, and by formal linguistic descriptions of cohesion, story grammars, speech acts, etc. Points at which this assumption, that adult intuitions of what makes a text more natural are appropriate for children, does not hold true should be identifiable from evidence that children have difficulty in decoding and/or comprehending adult constructed texts.

Expectations generally facilitate a reader's recognition of what comes next in a text by means of top-down processes that make some of the possible alternatives more likely than others, thereby also reducing the amount of further processing that is needed for positive identification of the new material, if it is consistent with what the reader expects. The reader's expectations can perform this function because they are based on knowledge which represents, to some degree, the regularities in the reader's experience with the linguistic and nonlinguistic world. Occasionally, however, expectations are violated by what actually comes next. Erroneous expectations reveal themselves most obviously when they result in overt reading miscues, but sometimes they may signal their presence in more subtle ways, such as hesitations, and at other times their immediate effects on the reading process may remain imperceptible to an observer. Some of the overt miscues resulting from erroneous expectations are of little or no consequence, as they do not change the essential meaning of the text. In those cases , it is as if the reader is engaged, probably unwittingly, in "copy editing" while reading. Other miscues are more serious because, if left uncorrected, they may affect immediate comprehension and perhaps compound themselves as the reader progresses further in the text.

The nature of a reader's expectations can vary, in part, as a function of the kind of text the reader thinks he is reading. For example, if the reader classifies a text as a story, his expectations will be different than if he sees the text as a set of instructions. Another thing that can vary, both within and between readers, is the extent to which expectations are employed at all. Where expectations based on linguistic and world knowledge have little influence, word recognition must of necessity be the product of mainly bottom-up processing of graphic cues. This might happen if the reader's expectations are violated too frequently, especially if they result in more serious word recognition errors of the sort that are unacceptable in the sentence or story context. It might also happen if the conditions needed for the formation and use of expectations are not met, as when a text contains too

many words that the reader cannot decode.

The preceding discussion suggests three basic strategies that children might use when reading: (1) a top-down strategy employing expectations based on natural written language; (2) a top-down strategy employing expectations based on primerese (which assumes that the reader has acquired some knowledge about the distinctive features of primerese and has classified the situation as one where primerese is likely to be employed); and (3) a bottom-up strategy employing graphic cues from the target word, together with knowledge about relationships between such cues and spoken words or word parts. (It might be said that this last strategy is based on a different kind of expectation, i.e., the expectation that close attention to the graphic cues in each word is needed in order to guarantee accurate word recognition.)

These three strategies do not exist in pure form, but the combination strategies that beginning readers actually use probably varyWor according to which strategy is emphasized. The particular kind of strategy that is emphasized in a given instance will vary as a function of the text and the surrounding context, as well as characteristics of the reader, and once the reader begins reading, it may be modified according to the reader's evaluation of the current strategy's effectiveness.

It seems unlikely that reader expectations and top-down processes will be emphasized unless the reader can recognize a fair number of words in the text with sufficient ease to read with some degree of fluency. In order to meet this condition in the present study, we screened the first graders who were potential subjects on their ability to recognize individual words from the stories to be read. Among those who passed the minimum criterion on our word recognition test, there was still a considerable range in word recognition ability. Consequently we divided the selected subjects into groups with high and low word recognition scores, so that we could determine whether recognition level had an effect on any differences we found between the original and rewritten versions.

METHOD

Design

The study employed a 2 x 2 matched pairs design in which 64 first grade students were matched on the basis of word recognition scores and randomly assigned to read one of two versions of four stories. The stories were either original versions of basal stories or rewritten versions of the same story in which specific features of primerese had been eliminated. The independent variables were story version (original or rewritten) and subject word recognition level (high or low). The dependent variables were scores on a set of comprehension questions, and several measures of oral reading performance. The basic design was replicated in each classroom involved in the study, in order to control the effects of classroom differences.

Subjects

Sixty four subjects were selected from a population of white, lower middle class first grade students in five classrooms in two communities. The classes were using basal readers as part of their normal reading instruction. Approximately two months before collecting the reading data, we asked the teachers to read trade books to the children on a regular basis, as well, to insure that the children all had some exposure to more natural written language.

Children were selected for the study if they reached a criterion of 30 words correctly read on the word recognition test. They were tested in January or February and then retested in March or April if they had not reached criterion on the first testing. Those Children who did not reach the criterion were excluded from the study. When children passed the criterion, they were then matched on the basis of word recognition test scores with other children from the same classroom and each member of the matched pair was randomly assigned to either the original or the rewritten set of stories. For purposes of analysis, the children in each text condition were further divided at the overall median into low (scores of 30 to 47) and high (47 or more) word recognition groups.

Procedures

The word recognition test was administered individually in a separate session. Subjects read the word list until they missed 5 words in a row, at which time sequential reading was discontinued and they were asked if there were any words remaining that they could read. Each session was tape recorded.

After assignment to the original or rewritten condition, subjects read four stories in separate sessions over a two week period. The order of the stories was the same for all subjects. Subjects were tested individually and the sessions were tape recorded. In each session, subjects were asked to read the story aloud two times. First subjects read the story title and, if they could not get a word, were given it. The researcher then asked a "motivating question", like "Read to find out how the hole gets fixed". The researcher asked for the answer to the motivating question between the first and second readings as a preliminary test of comprehension.

The reading was not interrupted by the researcher, except in cases where a subject skipped a line or more of text. If a subject paused for 5 seconds or more the researcher provided (pronounced) the word. When subjects read the story a second time, they were told that the researcher would not help them with words, and that they should try to read them or else skip them and go on.

After the second reading, subjects were asked 10 comprehension questions about the story.

Stories

The selections from the first grade texts in several reading series were examined. We were looking for selections that were already stories or could be made into stories with a little modification. The word "story" is used here in a minimal sense as possessing the basic elements of a beginning, middle, and end, a protagonist who tries to achieve a goal, etc. We found that a large number of selections in first grade readers are not stories in this sense. From a very limited set of alternatives, we chose three fables and one original story that would be read towards the middle of the year or later. The three fables were the tortoise and the hare, which was called "Rabbit and Turtle" (Clymer and Parr, 1976), "Lion and Mouse" (Clymer and Parr, 1976), and the fisherman and his three wishes, which was called "Bad Wish" (Raskin, 1984). The fourth story was about a truck and a bump in a road, and people's efforts to fix the road; it was called "Look Out" (Hershman, 1984). (See Appendix A).

The primerese or original versions of the texts were the selected stories left intact and presented in the readers with the accompanying illustrations. The rewritten versions were pasted over the original texts in the readers, so that they could be presented with same accompanying illustrations.

In order to make the language of the stories more like natural written language, and to make the content of the text clearer and more explicit, the rewriting included the following kinds of changes.

1. Increasing narrative structure.

This was accomplished by adding conventional openings and closings, as well as transitions between episodes, by changing dialogue carriers into their more common written forms, by changing dialogue to narrative description where the latter seemed more appropriate, and by changing the order of sentences where the original order had an illogical sequence.

In (1) a conventional opening has been added to the first sentence in the story, while in (2) a closing describing the condition of the two major characters at the end of the story has been added in the rewritten version. Additionally, dialogue has been changed to narrative description.

(1)
Original

The people came to the road.
Then a fast truck came.

Rewritten

One day a fast truck came down a road.

(2)
Original

He can't have his ship.
He can't have his help.
His wish was too big.
It's not good to stay.
I will not stay. [end of story]

Rewritten

"And now he can't have
his big ship or his helpers."
So the man and the woman
were back in the small boat. [end
of story]

(3)
Original

We will run to the park.
Rabbit said, "I want to stop here.

Rewritten

Let's run to the park.
We'll see who get's there first.
So Rabbit and Turtle began to run.
Soon Rabbit stopped to rest.

In (3) the original version does not state explicitly that Rabbit and Turtle began the race; instead, the scene shifts abruptly to the park, where Rabbit stops to rest. In the rewritten version an explicit transition has been added stating that the race had begun.

(4)
Original

"We have help on the ship," the woman said.
"Let the fish go."

Rewritten

"Now we have helpers," said the woman.
"So we can let the fish go."

In (4) The dialogue carrier in the original version is changed from a word order that is more typical of oral language, "the woman said," to a form in the rewritten version that is more commonly found in written language, "said the woman." In the original version the dialogue carrier could be misread as being attached to the following sentence if the punctuation is ignored. In the rewritten version the word order of the dialogue carrier makes this miscue

less likely because it would be ungrammatical.

(5)
Original

"Yes, I did," said Mouse.
"But you did not eat me, Lion.
"You did something for me, too."

Rewritten

"Yes, I did," said Mouse.
"But you did something for me, too.
You didn't eat me."

Finally in (5) the original version was been changed to move from the general statement to the specific example, which seemed more natural.

2. Increasing cohesion.

Cohesion was increased by pronominalizing repeated noun phrases, by combining sentences, and by marking relationships between sentences and clauses with conjunctions such as but, and, then, and so.

(6)
Original

"This road is not good.
This road has a bump.
This road has a hole ,too.

Rewritten

"This road is no good.
It has a bump
and it has a hole, too.

In (6) the noun phrase "This road" is repeated in three consecutive sentences in the original version. In the rewritten version the noun phrase is pronominalized and the second and third sentences are combined as two independent clauses by the addition of a conjunction.

(7)
Original

"I can't run fast.
You can't run, Turtle.
You can't run fast."

Rewritten

"I can run fast, but you can't.

In (7) more cohesion is created by combining the two sentences into one sentence with two clauses, and the explicit contrastive relationship is marked by the "but" in the rewritten version. Also the second run is ellipted. The third redundant sentence is dropped.

3. The elimination of unmotivated repetition and unnatural wording.

The elimination of repetition, a ubiquitous feature of primerese, often involved the combining of sentences. This can be seen in (8) where unmotivated repetitions of "the bump' and "the truck" were eliminated and sentences were combined. The odd locution "was on the bump" was changed to the more normal "hit the bump."

(8)
Original

The truck was on the bump.
The bump made the truck go up.

Rewritten

The truck hit the bump
and up it went.

In (9) the excessive repetition of "stop!" is deleted. Both instances of "Go slow" are deleted in (10) and replaced with the more natural wording "Slow down".

(9)
Original

"Stop!" the people said.
"Stop! Stop!

Rewritten

"Stop!" said the people.

(10)
Original

"Go slow, truck! Go slow!"

Rewritten

"Slow down!"

In (11) the same idea concerning the evaluation of the wish is repeated in different words in the original. The first statement is a circumlocution. In the rewritten version one sentence expresses the idea directly.

(11)
Original

It's not a good wish.
It's a bad wish.

Rewritten

It's a bad wish.

In (12) the verb form "got out" is used in the original version in an odd way, because it usually takes an animate subject. The sentence suggests that the truck got itself out. In the rewritten version the ambiguity is cleared up by making "the man" subject of the verb. The unnatural way of describing a bump that remains as being "up" is replaced with the more natural way of describing it as "still there."

(12)
Original

The truck got out.
But the hole was still there.
The bump was still up.

Rewritten

The man got his truck out
but the hole was still there
and so was the bump.

In (13) the circumlocution is replaced by the more natural wording "Let's" and the repetition is eliminated.

(13)
Original

You and I will run.
We will run to the park.

Rewritten

Let's run to the park.

4. Making explicit in the text information that is implied or carried in the illustrations.

In (14) the original version never explicitly states that the man got his big ship. It must be inferred. In the rewritten version this is stated explicitly.

(14)
Original

The fish said, "You may have your wish.
You may have a big ship. See!"
"I like my ship," the man said.

Rewritten

"You may have your wish," said the fish.
So the man got a big ship, and he was happy.

(15)
Original

Lion said, "Help! Help!
I want to get away from here, but I can't.

Rewritten

When Mouse met Lion again,
Lion was in a net and he could not get out.

After Lion has released Mouse the next episode begins on another page with a illustration of Lion in a net and Mouse on top of it. The Lion's problem is shown in the illustration and the text in (15) refers to it with the exophoric adverb "here," which is dependent upon the illustration for its interpretation. The rewritten version puts explicitly in the text the information carried in the illustration and, in doing so, eliminates the exophoric reference.

The overall results of the rewriting, in quantitative terms, can be seen in Table 1. As a result of the rewriting, the total number of words increased for two of the stories but not for the other two. In one case (LO) the story was actually made shorter. Primerese contains a great deal of repetition and redundancy, and the elimination of it therefore reduces the length of the stories. However, this reduction tends to be offset by the words that are added to make the text more explicit. Thus some of the rewriting changes shorten the text while others lengthen it.

The number of different words is generally greater in the rewritten versions than in the original primerese versions. Story number three (LO) is an exception. The loosening of controlled vocabulary in the rewritten versions would be expected to produce a greater variety of words. There is more repetition of words in the original versions as can be seen in the figures for the number of words per different words.

There are more sentences in the original stories and consequently the mean sentence length is greater in the rewritten versions, even where the original version contains more words (LO). The addition of explicit connectors and the combining of sentences increases mean sentence length. A readability analysis using the Fry (1968) formula shows that both versions are at first

Table 1

Number of Words, Different Words, and Sentences in Original (O) and Rewritten (R) Versions of Stories

Story	Version	Total Words	Different Words	Tot. Wds./ Diff. Wds.	Sentences	Words/ Sentence
RT	O	121	30	4.0	26	4.7
	R	135	57	2.4	19	7.1
LM	O	195	50	3.9	38	5.1
	R	215	69	3.1	35	6.1
LO	O	187	63	3.0	41	4.6
	R	161	63	2.6	25	6.4
BW	O	248	50	5.0	52	4.8
	R	247	66	3.7	45	5.5

grade level, but the rewritten versions are closer to second grade level on the graphs, indicating that there is some increase in difficulty according to traditional notions of readability.

Word Recognition Test

The word recognition test consisted of 65 words graded in difficulty from short, monosyllabic, frequently occurring words to less frequent polysyllabic words. It included most of the content words common to all the stories plus selected words that were unique to specific stories. Most of the new words that were added to the rewritten versions of the stories were also included. (See Appendix B).

Comprehension Test

The comprehension test consisted of 11 questions for each story. Most of the questions were factual, and some of these asked about information that was carried only in the illustrations in the original version and in the text and illustrations in the rewritten version. Other questions for each story were more general and required evaluation of behavior (See Appendix A). Some examples of factual questions were: Where did the animals decide to run to? What did Mouse do to help Lion?. Some examples of more general evaluative questions were: Why did Mouse help Lion? How did they feel when she fixed the road? Why was the man's fish bad?

Reading Miscue Coding

The coding of reading miscues involved a two stage process. In the first stage the tapes of the story readings were transcribed. All deviations of the observed oral reading responses from the expected responses were transcribed onto copies of the stories, using the transcribing system developed by Simons and Chambers (1981).

In the second stage the miscues from the transcriptions were coded into miscue categories which included the following error types: substitutions, omissions, insertions, nonsense words, partial word responses, words pronounced by the tester, words decoded by sounding out, pauses, and punctuation miscues. In addition, substitutions, omissions, insertions, and nonsense words were further coded for the features of syntactically acceptability, meaning change, graphic similarity, and self correction. A miscue was scored as syntactic acceptable if it was grammatically acceptable with the preceding context. Meaning change had two categories minor or major meaning change and was adapted from a system developed by M. S. Ammon (see Wong Fillmore et al., 1983) which was similar to Goodman's. A minor meaning change consisted of a miscues that only had a minor effect on the meaning of the original sentence. Minor changes included paraphrases, changes in tense or number, and function word substitutions. Major meaning changes included syntactically acceptable miscues which made the meaning ambiguous or confused and syntactically and semantically deviant miscues which made the sentences deviant and confused in a major way. A miscue was coded as graphically

similar if the first letter of the observed response in written form was the same as that of the printed word or, in the case of words of two or three letters, if the last part of the word was the same. All other responses were coded as graphically nonsimilar. Omissions were coded as graphically nonsimilar and insertions were not coded on graphic similarity. If the reader spontaneously corrected a miscue without intervention from the tester it was scored as self corrected.

Two other major reading performance variables were reading rate, which was the number of words read per minute, and error rate, which was the number of substitution, omission, insertion, and nonsense responses per 100 words. Finally, punctuation miscues included any significant deviation in intonation from what the punctuation predicted.

Each story was transcribed and coded twice. Transcribing involved noting any deviations from the expected response. Coding involved placing the deviations into the miscue categories. Both the second transcribing and coding served to correct any mistakes in the first transcribing and coding. As a reliability check, three subjects were independently coded by two coders and no significant discrepancies were found.

Approach to Quantitative Analyses

Although a matched-pairs design was employed in assigning subjects to text versions, the results reported below are not based on statistical procedures in which matched pairs were treated as the units of analysis. Two principal considerations made the use of such procedures undesirable. First, it was not always possible to match all pairs as closely as we wished within a given classroom. Second, there were several instances in which some data were missing for one or the other member of a particular pair. In some cases, these missing data resulted from such problems as failure to obtain a usable recording of the reading session in question. More commonly, however, the data were missing because the subject had gone through a particular reading of a particular story without committing any miscues. Where oral reading variables were defined as a proportions of total miscues, we could not assign the subject any scores for those variables. Had we used statistical procedures appropriate for matched pairs, the occurrence of missing data for one member of a pair would have made it necessary for us to omit the other pair member from the analysis in question, even when relevant data from the second member were available. Thus, in order to use all of the available data, we have treated text version as a between-subject factor that is crossed with or nested within word recognition level. In a 2x2 factorial design for the analyses of variance we will report below.

The occurrence of missing data is also problematic for analyses of such within-subject factors as differences between stories, or between first and second readings of the same story. Statistical tests of these factors are not of primary importance with regard to our research questions, however. Thus in the quantitative analyses of oral reading performance we will focus on separate 2x2 analyses of variance for each reading of each story. Because the overall number of these analyses is quite large--eight ANOVA's for each of several dependent variables--we will be concerned principally with those

effects that were statistically significant in at least two of the four stories. Moreover, in interpreting the results, we will emphasize the overall pattern of findings, rather than any single finding. Most of the analyses of variance we performed did not reveal significant interactions of text version by recognition level. Consequently the results of these analyses will be reported primarily in terms of the main effects of those two factors. Interaction effects will be displayed only where they are statistically significant.

Results

Results will be reported first for the children's responses to the comprehension questions and then for their oral readings, which were analyzed both quantitatively and qualitatively.

Comprehension Questions

Each student received a comprehension score equal to the number of comprehension questions answered correctly, out of the eleven that were asked for each story. A total comprehension score was obtained by summing across the four stories. Means of all of these scores are displayed in Table 2 by text version and word recognition level. Performance on the comprehension questions was generally quite high, with the average subject answering over 70% of all questions correctly, regardless of recognition level or text version. Text version appears to have made virtually no difference within the low recognition group, but there were consistent differences favoring the rewritten versions within the high recognition group, as statistically significant effects of text version occurred in two of the stories and in total score.

It is interesting to note that the high recognition subjects reading the original version did no better than the low recognition subjects reading the same version, despite their greater ability to recognize individual words. Apparently the low recognition subjects could recognize enough of the words to comprehend the same amount of the content in the original stories, at least after two readings. However, the high recognition subjects were able to profit from the changes made in the rewritten version, whereas the low recognition subjects - on balance - were not helped very much, but clearly were not harmed either.

An analysis of the individual comprehension questions revealed that a majority of the questions about each story were answered correctly more often by students who read the rewritten version than by those who read the original. Overall, 32 out of 44 questions favored the rewritten version. A series of exact probability tests showed that 7 of those 32 items discriminated significantly between the two text versions, and there were no significant differences favoring the original versions. The number and size of the differences favoring the rewritten versions may have been greater had there not been a tendency toward a ceiling effect on several of the questions. That is, the questions were answered correctly by so many subjects in both text conditions that there was little chance of finding group differences on those items.

Table 2

Mean Comprehension Scores for Original and Rewritten Version Within Low and High Word Recognition Groups

	Low Recognition		High Recognition	
Story	Original	Rewritten	Original	Rewritten
RT	9.0	9.1	7.7	9.1[a]
LM	6.8	7.4	6.6	8.4[b]
LO	7.9	8.8	8.4	9.3
BW	7.0	6.9	7.0	8.1
Total	30.2	31.6	30.1	34.9[c]

[a] $t(26) = 2.15$, $p < .05$.
[b] $t(26) = 2.79$, $p < .01$.
[c] $t(26) = 2.49$, $p < .05$.

The 7 questions on which the revised version led to significantly more correct responses suggest some reasons for the overall superiority of that version. In four cases, a correct answer required information that was stated explicitly in the rewritten version but was only carried by pictures in the original version, or was implicit in the text. For example, in the original version of "Lion and Mouse" Mouse says:

"Look here Lion.
See what I can do."

and an illustration shows Mouse nibbling a hole in the net in which Lion is caught. The rewritten version reads as follows:

"Look here, Lion," said Mouse.
"Look what I can do."
Mouse nibbled a hole in the net
and Lion got out.

When asked "What did Mouse do to help Lion?," students who read the rewritten version frequently used the word "nibbled" in their answers, while those who read the original version did not, even when they answered correctly. Thus it appears that students reading the rewritten version obtained the necessary information from the text, and not from the pictures.

There were two other questions that required information that was implicit in both the original and rewritten version. The fact that more correct responses occurred with the rewritten version suggests that the language of the version made it easier for readers to coordinate different parts of the text and make the inferences which were needed to answer these questions. Finally, there was one question, "what were the three wishes that the man made?," which required information that was explicit in both versions. Once again, however, the reader had to coordinate different parts of the text in order to answer correctly, and the rewritten version apparently made this task easier.

These findings suggest that the rewritten version may have allowed children to make greater use of expectations based on natural language in top-down processing of the texts. To pursue this possibility further, we turn now to a more direct look at the reading process, through our analyses of the children's oral reading behavior.

Quantitative Analyses of Oral Reading

With regard to these quantitative analyses, we will generally report the main effects of word recognition level first, to establish whether the dependent variables of interest are sensitive to an independent variable that ought to affect reading, i.e., the child's ability to recognize individual words contained in the stories. With these results as background, we will then turn to the results that are of principal interest, the main effects of text version, and the interaction of text version and recognition level that were statistically significant.

Reading rate and miscue rate. Two global measures of reading difficulty were derived from the subjects' tape recorded oral readings. The first was reading rate, expressed as the number of words read per minute. The second measure, miscue rate, was obtained by summing instances of substitution, omission, insertion, and nonsense responses to obtain the total number of miscues per 100 words. These are "word" miscues, i.e., miscues in which the subject produced a word or word-like form that was different from the one on the printed page. The means displayed in Table 3 show the main effect of word recognition level on both of these dependent variables. It is clear that both measures discriminated strongly and consistently between subjects with high and low word recognition scores. That is, the high recognition group read at a faster rate and with fewer miscues than the low recognition group.

Main effects of text version on the same two variables are shown in Table 4. With regard to reading rate, there was a general tendency for the rewritten versions of the stories to be read at slightly slower rates than the original versions, but none of these version differences were statistically significant. With regard to overall miscue rate, significantly more miscues occurred on the rewritten version in both readings of "Rabbit and Turtle", but no significant differences were observed for any other story. Thus there does not seem to have been a general tendency for subjects reading the rewritten version to have been more miscue-prone. However, text version did interact significantly with recognition level for the first readings of both "Lion and Mouse" and "Bad Wish." The means for these interactions are displayed in Table 5. They show that, in comparison with the subjects reading the original version, the miscue rates on the rewritten version were higher for low recognition subjects but lower for high recognition subjects. Thus, with the rewritten versions of these two stories, the general tendency toward more miscues on the part of low recognition subjects was amplified.

The data in Tables 4 and 5 show that, in some instances, the rewritten versions resulted in higher rates of miscues than the originals, i.e., in "Rabbit and Turtle" and, for subjects with low word recognition scores, in "Lion and Mouse" and "Bad Wish" as well. However, rather than move directly to the question of why the miscue rates may have increased in these instances and not in others, we will turn instead to some more finegrained miscue analyses by examining the frequency with which different kinds of miscues occurred, in relation to both word recognition level and text version. Some miscues might be regarded as "better" or "worse" than others, or as indicative of one approach to reading versus another. Consequently, differences in total miscue rate are difficult to interpret at face value, and similar miscue rates may mask important differences in the ways different versions of a text are read. Likewise, similarities in overall reading rates, such as those shown in Table 3, may mask important differences in the flow of oral reading over time.

Substitutions, omissions, insertions, and nonsense responses. One way to accomplish a more finegrained analysis is by examining each of the various miscue categories that constituted the total miscue score, i.e., substitution, omission, insertion, and nonsense miscues. The proportions of total miscues occurring in each of these categories were subjected to the same sorts of analyses of variance as those already described above. Means for the main

Table 3

Mean Reading Rates and Miscue Rates for Students

With Low (L) and High (H) Word Recognition Scores

	Reading Rate[a]		Miscue Rate[b]	
Reading	L	H	L	H
RT 1	46	81	3.6	1.7
2	57	91	3.7	1.5
LM 1	47	83	5.2	2.8
2	62	95	5.3	3.0
LO 1	49	89	4.5	2.7
2	60	106	4.1	2.5
BW 1	47	83	6.9	3.8
2	56	99	6.5	3.8

[a] Number of words read per minute. All differences between L and H significant at $p<.001$.

[b] Number of miscues per 100 words. All differences between L and H significant at $p<.01$, except for BW 2, where $p<.05$.

Table 4

Mean Reading Rates and Miscue Rates for Original (O) and Rewritten (R) Versions

	Reading Rate[a]		Miscue Rate[b]	
Reading	O	R	O	R
RT 1	67	62	1.7	3.6[c]
2	76	75	1.9	3.4[d]
LM 1	68	68	3.7	4.3
2	82	80	3.9	4.4
LO 1	74	70	4.1	3.1
2	87	85	3.6	3.1
BW 1	69	65	5.0	5.6
2	83	77	4.6	5.3

[a] Number of words read per minute.
[b] Number of miscues per 100 words.
[c] $F(1,56) = 12.24$, $p < .001$.
[d] $F(1,56) = 7.26$, $p < .01$.

Table 5

Mean Miscue Rates in Significant Interactions of Text Version and Word Recognition Level

Reading	Recognition Level	Version	
		Original	Rewritten
LM 1	Low	3.9	6.7[a]
	High	3.5	2.2
BW 1	Low	5.5	8.1[b]
	High	4.6	3.0

[a] Interaction $F(1,52) = 5.28$, $p<.05$.

[b] Interaction $F(1,52) = 4.80$, $p<.05$.

effects of word recognition level on these proportions are displayed in Table 6. Regardless of recognition level, a large majority of the miscues students made were substitutions. There was a consistent tendency for subjects with low recognition scores to make more substitution miscues than those with high recognition scores, but this trend reached significance for just one reading of one story. Clearer differences between recognition levels are apparent with respect to omission miscues, as high recognition subjects produced significantly greater proportions of omissions in three of the four stories. The proportions of insertions and nonsense miscues were generally very low, and there was only one instance of a significant difference between recognition levels, in which a higher proportion of insertions was found for high recognition subjects.

To the extent that the differences shown in Table 6 are reliable, they could simply reflect the fact that subjects with higher recognition scores generally read at faster rates, and may therefore have been more likely to omit a word now and then. But these data might also indicate that subjects in the high recognition group were engaged in more top-down processing of the sort that entails coordination of the preceding and following context when reading a particular word, and that they were not simply reading one word at a time. A similar difference between original and rewritten versions would therefore be interesting, as it would suggest that one version is more conducive to such processing than the other. However, the means in Table 7 for the main effects of text version provide virtually no evidence of systematic differences between the original and rewritten version with respect to the proportions of miscues that occurred in the various categories. In only two isolated instances were version differences statistically significant (see Table 7) and in no case did version interact significantly with recognition level.

Syntactic, semantic, and graphic miscues. The categories used above provide more information than the overall miscue score, but they still suffer from some of the same limitations too. For example, different substitutions and omissions may vary in their acceptability, or they may reflect different reading processes. It seems important, therefore, to look at each miscue in relation to different sorts of cues provided by the text. As in past research, we here distinguish three sorts of text cues: syntactic, semantic, and graphic. That is, a given miscue may or may not fit the preceding syntactic context, it may or may not change meaning, and it may or may not reflect the graphic properties of the to-be-read word. Miscue types based on these distinctions are not entirely orthogonal to the miscue categories analyzed above, but they do provide different and potentially more revealing perspectives on the kinds of miscues readers make.

Table 8 contains means for the main effects of word recognition level on the proportions of miscues that reflected the syntactic, semantic, and graphic cues in the text being read. It can be seen that, in general, a majority of the miscues did fit the syntactic contexts in which they occurred, and that this happened proportionally more often for subjects with high word recognition scores; significant differences between the high and low recognition groups were found in three of the four stories. Miscues that fit the semantic context were somewhat less frequent--generally just under 50%.

Table 6

Mean Proportions of Substitution, Omission, Insertion, and Nonsense Miscues for Students with Low (L) and High (H) Word Recognition Scores

	Substitution		Omission		Insertion		Nonsense	
Reading	L	H	L	H	L	H	L	H
RT 1	.81	.69	.06	.17[a]	.04	.06	.09	.08
2	.80	.74	.06	.17[b]	.04	.03	.10	.06
LM 1	.79	.70	.05	.11	.06	.16[b]	.10	.03
2	.75	.71	.07	.14	.07	.10	.12	.06
LO 1	.82	.74	.07	.16[c]	.02	.06	.09	.04
2	.84[d]	.69	.06	.15[e]	.05	.12	.05	.04
BW 1	.86	.77	.04	.11[f]	.03	.08	.07	.05
2	.84	.71	.02	.20[g]	.08	.07	.05	.03

[a] $F(1,51) = 5.33$, $p<.05$.
[b] $F(1,51) = 4.12$, $p<.05$.
[c] $F(1,50) = 4.62$, $p<.05$.
[c] $F(1,55) = 4.73$, $p<.05$.
[d] $F(1,55) = 5.84$, $p<.05$.
[e] $F(1,55) = 4.16$, $p<.05$.
[f] $F(1,52) = 5.06$, $p<.05$.
[g] $F(1,51) = 8.97$, $p<.05$.

Table 7

Mean Proportions of Substitution, Omission, Insertion, and Nonsense Miscues for Original (O) and Rewritten (R) Versions

	Substitution		Omission		Insertion		Nonsense	
Reading	O	R	O	R	O	R	O	R
RT 1	.74	.77	.10	.12	.09	.02[a]	.08	.10
2	.79	.76	.12	.10	.04	.03	.04	.11
LM 1	.76	.73	.08	.09	.09	.13	.07	.06
2	.73	.73	.12	.09	.08	.08	.08	.10
LO 1	.82	.74	.07	.16[b]	.03	.05	.08	.04
2	.77	.76	.11	.10	.08	.09	.04	.04
BW 1	.83	.80	.07	.13	.06	.05	.04	.08
2	.73	.82	.14	.10	.09	.06	.04	.04

[a] $F(1,56) = 4.51$, $p < .05$.

[b] $F(1,55) = 5.78$, $p < .05$.

Table 8

Mean Proportions of Syntactic, Semantic, and Graphic Miscues for Students with Low (L) and High (H) Word Recognition Scores

	Syntactic		Semantic		Graphic	
Reading	L	H	L	H	L	H
RT 1	.55	.65	.39	.44	.79[a]	.59
2	.65	.71	.42	.50	.66	.60
LM 1	.49	.69[b]	.42	.46	.75	.62
2	.52	.60	.43	.45	.65	.61
LO 1	.62	.64	.42	.44	.61	.54
2	.61	.77[c]	.42	.51	.64[d]	.45
BW 1	.50	.70[e]	.46	.49	.68[f]	.54
2	.51	.65[g]	.40	.55[h]	.61	.54

[a] $F(1,51) = 16.51$, $p<.001$.
[b] $F(1,51) = 10.88$, $p<.01$.
[c] $F(1,55) = 6.01$, $p<.05$.
[d] $F(1,53) = 8.83$, $p<.01$.
[e] $F(1,52) = 14.43$, $p<.001$.
[f] $F(1,52) = 4.55$, $p<.05$.
[g] $F(1,51) = 4.63$, $p<.05$.
[h] $F(1,49) = 4.32$, $p<.05$.

The high recognition group consistently attained higher proportions on this variable, but most of the differences between recognition levels were quite small, and only one of them reached statistical significance. Finally, while most miscues tended to be consistent with graphic cues contained in the target word, the proportion of such miscues was consistently higher for the low recognition group than for the high recognition group, with significant differences of this sort occurring in three of the stories.

Taken together, the results displayed in Table 8 provide clearer evidence than before that subjects with higher word recognition scores were engaged in more top-down processing, and were less likely to rely heavily on graphic cues in each succeeding word as they read. The principal question of interest, then, is whether these apparently more discriminating variables also show differences between the original and rewritten versions of the stories.

The means in Table 9 are for main effects of text version on the proportions of miscues fitting syntactic, semantic, and graphic cues in the text. With regard to miscues that were syntactically appropriate, two stories showed significantly higher proportions for the rewritten version, one showed higher proportions for the original version, and one showed virtually no difference between versions. In contrast with this rather mixed outcome, the version differences for miscues that were semantically appropriate all favored the rewritten version, and were statistically significant in two of the four stories. The results for graphically appropriate miscues are, again, mixed. In one case, a significantly higher proportion of graphic miscues occurred for the original version (see LO1), but in another case there was a significant difference in the opposite direction (see RT1); the remaining differences were generally quite small and were not consistently in one direction or the other. However, there were also significant interactions of text version and recognition level in four of the eight analyses of graphic miscues. The means for these interactions are presented in Table 10. Three of the four interactions appear to be similar in form (RT2, LO1, and BW1): for high recognition subjects, the proportion of graphic miscues was lower with the rewritten version; for low recognition subjects, the difference between versions was either reversed or at least eliminated. The fourth significant interaction (RT1) was another one in which the versions differed for the high recognition subjects, but in this case, a higher proportion of graphic miscues occurred with the rewritten version.

By way of summarizing the results shown in Tables 8, 9, and 10, we might again ask to what extent the effects of text version paralleled those of recognition level. The results are clearest with respect to miscues that were appropriate semantically, as these occurred in higher proportions both for high recognition subjects and for subjects reading the rewritten versions. While these differences tended to be small and only reached statistical significance in a few instances, they were entirely consistent in direction. The results for syntactic miscues are slightly more complex. The higher proportions attained by subjects in the high recognition group were also found for those reading the rewritten versions of two of the stories, but on another story (RT) it was the subjects reading the original version who had higher proportions of syntactic miscues. Still more complex are the results for graphic miscues. In contrast to the main effects of recognition level, which

Table 9

Mean Proportions of Syntactic, Semantic, and Graphic Miscues for Original (O) and Rewritten (R) Versions

	Syntactic		Semantic		Graphic	
Reading	O	R	O	R	O	R
RT 1	.70[a]	.51	.36	.46	.62	.77[b]
2	.79[c]	.59	.44	.48	.62	.65
LM 1	.60	.59	.37	.52[d]	.71	.64
2	.58	.54	.38	.50	.62	.64
LO 1	.55	.70[e]	.34	.52[f]	.68[g]	.48
2	.67	.71	.38	.53	.55	.54
BW 1	.57	.63	.45	.51	.61	.61
2	.51	.65[h]	.41	.53	.52	.62

[a] $F(1,51) = 4.43$, $p<.05$.
[b] $F(1,51) = 11.52$, $p<.01$.
[c] $F(1,51) = 6.54$, $p<.05$.
[d] $F(1,51) = 4.03$, $p<.05$.
[e] $F(1,54) = 4.56$, $p<.05$.
[f] $F(1,55) = 5.47$, $p<.05$.
[g] $F(1,53) = 11.29$, $p<.01$.
[h] $F(1,51) = 4.67$, $p<.05$.

Table 10

Mean Proportion of Graphic Miscues in Significant Interactions of Text Version and Word Recognition Level

Reading	Recognition Level	Version	
		Original	Rewritten
RT 1	Low	.78	.79[a]
	High	.34	.75
RT 2	Low	.56	.74[b]
	High	.69	.53
LO 1	Low	.61	.61[c]
	High	.78	.36
BW 1	Low	.60	.74[d]
	High	.62	.46

[a] Interaction $F(1,51) = 10.63$, $p<.01$.
[b] Interaction $F(1,51) = 4.31$, $p<.05$.
[c] Interaction $F(1,53) = 11.73$, $p<.01$.
[d] Interaction $F(1,52) = 5.82$, $p<.05$.

consistently showed greater proportions of graphic miscues by low recognition subjects, only two main effects of text version were significant, and these were in opposite directions. However, text version was also found to interact significantly with recognition level in four of the eight readings, and three of the interactions--from three different stories--showed that subjects in the high recognition group had higher proportions of graphic miscues if they read the original versions, while those in the low recognition group tended to make more graphic miscues if they read the rewritten versions.

Thus there were several instances in the data where subjects reading the rewritten versions were more likely to make the kinds of miscues that were more frequently associated with higher word recognition scores--i.e., miscues that were consistent with syntactic or semantic cues in the text, but not with graphic cues in the target word. The only significant exceptions to these trends occurred in readings of the "Rabbit and Turtle" story, where the usual text version effects were reversed for both syntactic and graphic miscues. We might recall, also, that "Rabbit and Turtle" was the only story to show significant differences between versions in overall miscue rate, and that the higher miscue rates occurred with the rewritten version.

Miscues on new words in rewritten versions. In order to enhance our understanding of the findings reported thus far--including the discrepant results from "Rabbit and Turtle"--we conducted further analyses of data from the rewritten versions in which we distinguished between miscue scores based on all words and the same kinds of miscue scores based only on "new" words, i.e., words that were added to the text when the original was rewritten. (Almost none of the original word types were dropped from the revisions, so it was not meaningful to do similar analyses of the original versions, distinguishing words that were not in the rewritten versions.) Table 11 compares mean miscue rates for new words and for all words, for both low and high recognition subjects who read the rewritten versions. It is obvious that, in "Rabbit and Turtle," the, miscue rates for new words were higher than those for all words, especially for subjects with low word recognition scores. In fact, for both recognition levels, the new word miscue rates for "Rabbit and Turtle" are the highest ones in Table 11. In the other stories, the rate of miscues on new words was, if anything, lower than on all words, but the differences were quite small. Thus the significantly higher overall miscue rates for the rewritten version of "Rabbit and Turtle" are due primarily to miscues made on new words in the rewritten version.

Similar analyses comparing miscues on new words with those on all words were carried out for the proportions of syntactic, semantic, and graphic miscues made by low recognition subjects on the rewritten versions, as shown in Table 12. (Means for high recognition subjects are not included in this table because their low miscue rates resulted in proportions based on small numbers of subjects with very low miscue frequencies.) With respect to the kinds of miscues found earlier to occur in higher proportions with the rewritten versions--i.e., syntactic and semantic miscues--proportions for new words in Table 12 are substantially lower than those for all words (the only exception being syntactic miscues on L02). In other words, subjects reading the rewritten versions tended to make proportionately more syntactic and semantic miscues overall despite the kinds of miscues they tended to make on words that were new in the rewritten versions.

Table 11

Mean Miscue Rates on "New" Words and All Words in the Rewritten Version for Students with Low and High Word Recognition Scores

	Low Recognition		High Recognition	
Reading	New Words	All Words	New Words	All Words
RT 1	21.4	4.7	5.2	2.5
2	16.5	4.9	3.9	1.8
LM 1	7.5	6.7	1.6	2.2
2	5.7	6.0	1.6	3.0
LO 1	3.1	3.5	1.0	2.7
2	2.9	3.8	0.4	2.3
BW 1	7.4	8.1	0.9	3.0
2	6.0	7.7	1.0	3.1

Table 12

Mean Proportions of Syntactic, Semantic, and Graphic Miscues on "New" Words and All Words in the Rewritten Version for Students with Low Word Recognition Scores

	Syntactic		Semantic		Graphic	
Reading	New	All	New	All	New	All
RT 1	.26	.38	.23	.42	.66	.79
2	.35	.55	.20	.45	.79	.74
LM 1	.26	.48	.10	.49	.54	.71
2	.29	.49	.13	.47	.77	.70
LO 1	.39	.64	.15	.55	.56	.61
2	.68	.62	.19	.52	.93	.67
BW 1	.28	.52	.24	.45	.70	.74
2	.44	.61	.15	.43	.77	.68

As for the proportions of graphic miscues shown in Table 12, the principal regularity seems to be a reversal between the first and second readings of each story, in which the proportion of graphic miscues for new words is below the one for all words on the first reading, but above it on the second, due to a substantial increase in the proportion of graphic miscues for new words from the first to the second reading. These increases may help account for the finding that graphic miscues generally did not occur in higher proportions for the original versus the rewritten versions within the second readings. However, the data on graphic miscues in Table 12 appear to shed no new light on the significant main effects of text version (see Table 9) or on the interactions of text version with recognition level (see Table 10).

Phonic and punctuation miscues. In addition to the "word" miscues discussed so far, two other types of reading miscue were subjected to quantitative analyses like the ones we have already reported. The first type, called "phonic" miscues, includes instances in which the subject made some incomplete attempt to sound a word out, sometimes leading the experimenter to provide the correct word. The second type of miscue, called "punctuation" miscues, includes instances in which the prosodic features of the subject's oral reading did not fit the printed punctuation in the text, e.g., sentence-final intonation in the absence of a period, or the lack of such intonation where a period was present. Mean rates of these two types of miscues per 100 words are shown in Table 13 as a function of recognition level.

It can be seen in Table 13 that phonic miscues generally were quite infrequent, but--not surprisingly--were more frequent for subjects with low word recognition scores than for those with high recognition scores. In contrast, punctuation miscues were somewhat more common, but they occurred with about equal frequency for the low and high recognition groups. If anything, there was tendency for the high recognition subject to make more punctuation miscues, but none of the differences in the rate of such miscues were statistically significant. At this point, it may seem somewhat puzzling that high recognition subjects do not appear to have made fewer punctuation miscues than their low recognition counterparts, but the analyses reported below begin to suggest an explanation.

When phonic and punctuation miscues are analyzed for effects of text version, the results do not parallel those for effects of recognition level, as indicated by the data in Table 14. With just one exception (RT1), the original and rewritten versions did not differ significantly with regard to rates of phonic miscues, nor was there any tendency for the differences to fall in one direction or the other. These negative findings might result from the very low rates at which phonic miscues generally occurred. In contrast, punctuation miscues were found to occur at higher rates for the original versus the rewritten version in all eight readings, and statistically significant differences of this sort were found in all stories except "Rabbit and Turtle." Moreover, while none of the interactions between text version and recognition level were statistically significant, examination of the means for these interactions (not shown here) revealed an interesting descriptive result: in all eight readings, the highest rate of punctuation miscues was produced by the high recognition subjects who read the original version, as compared to the high recognition subjects who read the rewritten version or

Table 13

Mean Rates of Phonic and Punctuation Miscues for Students with Low (L) and High (H) Word Recognition Scores

	Phonic Miscues		Punctuation Miscues	
Reading	L	H	L	H
RT 1	1.1[a]	0.3	1.3	1.4
2	0.8[b]	0.2	1.1	1.4
LM 1	1.2[c]	0.3	1.2	1.4
2	0.7[d]	0.1	1.3	1.2
LO 1	1.0[e]	0.2	1.5	2.2
2	0.6[f]	0.0	1.4	1.9
BW 1	0.7[g]	0.2	2.6	2.6
2	0.4	0.2	2.1	2.4

a $F(1,56) = 11.18$, $p<.01$.
b $F(1,56) = 10.04$, $p<.01$.
c $F(1,52) = 22.74$, $p<.001$.
d $F(1,51) = 16.43$, $p<.001$.
e $F(1,56) = 19.76$, $p<.001$.
f $F(1,56) = 13.02$, $p<.001$.
g $F(1,53) = 8.49$, $p<.01$.

Table 14

Mean Rates of Phonic and Punctuation Miscues for Original (O) and Rewritten (R) Versions

	Phonic Miscues		Punctuation Miscues	
Reading	O	R	O	R
RT 1	0.4	1.0[a]	1.7	1.0
2	0.3	0.7	1.4	1.1
LM 1	0.6	0.8	1.7[b]	0.8
2	0.5	0.3	1.5	0.9
LO 1	0.7	0.5	2.5[c]	1.1
2	0.5	0.2	2.4[d]	0.9
BW 1	0.4	0.5	3.7[e]	1.5
2	0.2	0.5	3.4[f]	1.1

[a] $F(1,56) = 4.983$, $p<.05$.
[b] $F(1,51) = 7.578$, $p<.01$.
[c] $F(1,56) = 12.619$, $p<.001$.
[d] $F(1,56) = 13.399$, $p<.001$.
[e] $F(1,52) = 21.193$, $p<.001$.
[f] $F(1,51) = 29.243$, $p<.001$.

the low recognition subjects who read either version. On the other hand, in four of the eight readings, the lowest rate of punctuation miscues was produced by the high recognition subjects who read the rewritten version. In other words, the subjects who generally seemed more inclined to employ top-down processes in their reading (i.e., the subjects with higher word recognition scores) were especially likely to make punctuation miscues if they read the original versions, but were unlikely to make such miscues if they read the rewritten versions. This finding suggests that their top-down processes entailed expectations which were more often confirmed by the rewritten than by the original versions. It also explains why there were no main effects of recognition level such that high recognition subjects, on the whole, made significantly fewer punctuation miscues than low recognition subjects (see Table 13).

Summary of quantitative findings. The rewritten versions employed in this study appear to have had virtually no effect on the overall rate at which students read the stories, nor did they consistently affect the overall rate of word miscues (substitution, omissions, insertion, and nonsense), although evidence of significant increases in miscue rate was found for all subjects on one story ("Rabbit and Turtle") and for low recognition subjects on two other stories. The increase in "Rabbit and Turtle" can be attributed to unusually high miscue frequencies on some of the words that were "new" in the rewritten version. More important, however, the miscues made on the rewritten versions of all the stories were consistently more likely to be semantically appropriate, and - for children with higher word recognition scores - they were less likely to be based primarily on graphic cues in the word to be read. In addition, the rewritten versions consistently resulted in fewer punctuation miscues, particularly for children in the high recognition group. Taken together, these findings suggest that the rewritten versions encouraged greater use of top-down versus bottom-up processing, especially on the part of children who were already capable of recognizing a relatively large number of the individual words in the texts. Moreover, they did so with little or no decrement in the accuracy or rate of word recognition. These findings suggest that the changes made in the written versions brought the texts more in line with the kinds of expectations likely to be formed by young readers. The following qualitative analyses will explore this interpretation further.

Qualitative Analysis of Reading Miscues

Oral reading miscues provide evidence and insight into children's top-down and bottom-up reading strategies. As argued above, top-down strategies can involve expectations based on either natural language or primerese. Top-down miscues occur when readers have strong expectations that are in conflict with the text. These miscues provide insights into the effects of primerese features on children's reading.

The miscues produced by the conflict between natural language expectations and primerese features include both editing and repunctuation miscues. Editing miscues are substitutions, insertions, or omissions that make the primerese text more semantically explicit, more natural in form, or more independent of the illustrations (Simons & Elster, 1985). Repunctuation miscues (Chambers, 1987) involve deviations from the expected intonation

patterns in sentences, as indicated by the printed punctuation. In traditional reading error terminology they would be characterized as ignoring punctuation. Repunctuation miscues show readers expectations about sentence structure, the form and placement of dialogue carriers, and exclamation points and question marks.

With children who have had frequent exposure to primerese texts, reading miscues also show the effect of expectations about the nature of primerese. These expectations show up in what we call repeated structure miscues, which are characterized by the repetition of words or phrases from the preceding text. Such miscues appear to result from the excessive and unmotivated repetition found in primerese texts.

Bottom-up reading strategies are indicated by nonsense word substitutions, repeated attempts at a word, long pauses before words, and slow word by word reading with citation pronunciation. They may also be indicated by a relatively high graphic similarity scores for real-word substitutions. By definition, however, the miscues associated with more bottom up processing tell us little about the expectations involved in top-down processing except to suggest that those expectations would be at odds with the text at hand, which may be the reason why the reader is engaged in bottom-up processing. Therefore, this discussion will focus on top-down miscues and the information they provide about readers' expectations.

The miscues to be discussed in this section occured primarily in the original primerese versions. They also appeared, less frequently, in the rewritten versions. However, their occurrence in the rewritten versions was almost always at places in the texts where the rewritten version and the primerese version were the same.

In interpreting reading miscues it should be kept in mind that the printed text exerts the strongest influence by far on what is produced orally. There is a very strong tendency for readers when reading aloud to reproduce the printed text without any deviations. Simons and Leu (1987) have shown that subjects' reading of text that was disrupted by words which did not fit the context almost invariably read the printed words even though they produced anomalous sentences. This means that the vast majority of the words that are read in any text will not produce miscues. As we reported earlier, the overall rate of miscues in the present study was generally below 5 per 100 words. Even where the primerese texts used in this study conflict with children's expectations, we can only expect a few readers to produce a miscue, because of this tendency to reproduce the printed text. Moreover, miscues are self corrected a high proportion of the time, providing further evidence of the salience of the printed text. The structures may conflict with the expectations of the children who produce no miscues, but this conflict does not reveal itself in a miscue. This makes it difficult to draw inferences about the reading process on the basis of the non-occurrence of an error. It also makes quantification of miscues on specific features difficult because of their low frequency of occurrence.

Even when top-down miscues do occur they may also be heavily influenced by the graphic features of the printed text. They tend to maintain the first

letter and length of the printed word (Weber, 1970). This means that the influence of the graphic features of the text must often be taken into consideration, in addition to structural features, in interpreting miscues. Miscues appear in many cases to be influenced both by the graphic features and by structural expectations. This analysis will focus on structural expectations, but the graphic features need to be kept in mind as alternate or joint explanations of the miscues. As Goodman (1973) pointed out, miscues do provide a window on the reading process. This window is not as clear and unambiguous as we would like it to be, but we think the insights to be gained from the study of reading miscues far outweigh the problems of interpretation.

Natural Language Expectations

Editing Miscues

Editing miscues, which occur in the original versions but rarely in the rewritten, make the text conform to the readers natural language expectations. These miscues have three major related effects: they make the semantic content of the text more explicit, they translate odd or awkward structures into more natural written prose, and they put information in the text that is only carried in the illustrations accompanying the text.

Making content explicit. Primerese, as we have noted, often leaves the semantic content implicit rather than stating it explicitly in the text. This feature requires more inferencing on the part of the reader. The editing miscues sometimes put into words semantic content that is implicit. Some examples can be seen in (16) and (17).

(16) I can get ahead ^ to the park faster first. fast.

(17) You and I will race run. We will race run to the park.

The content of the Rabbit and Turtle story concerns a race and the relationship of the speed of the two contestants. The text is inexplicit with regard to these issues and does not use the words "race", or "first" or the comparative "faster," which would normally be used in this story. In the above examples, the readers' miscues edit the text to make the language more explicit with respect to a race schema. In (16) readers substituted the word "first" or "faster" for "fast," or inserted the word "ahead" to make the sentence state more explicitly that one of the contestants would reach the finish ahead of the other. One subject actually substituted the word "race" for "run," as can be seen in (17). This substitution represents the race schema explicitly in the text. This last miscue might be seen not as a schema-based miscue, but as a single word substitution in which the subject uses a near synonym for "run" that also begins with the letter "r". However, if it is a schema- based miscue, which seems plausible in light of the other miscues shown above, it suggests a rather high level of inferencing.

In (18) the contrast between the speed of Rabbit and Turtle is made more explicit by the insertion of the contrastive conjunction "but."

(18) You run fast Rabbit ^ (but) I can't run fast.

In (19) and (20), which come from the rewritten version, the conjunction "and," which does not explicitly make the comparison, is deleted. The deletion of "and" makes the comparison a little more explicit because it removes a coordinating conjunction. The contrast is made fully explicit in (20), which contains the contrastive conjunction "but".

(19) I can run fast, (and) you can't.

(20) Turtle said, "You can run fast, Rabbit, (but) and I can't

(21) But you stopped and I didn't. That's (how) why I got here first.

In (22) the substitution of "how" for "why" again makes the meaning more explicit. The latter part of the first sentence states what Turtle did to arrive first, i.e. he did not stop as Rabbit did. The "how" substitution miscue refers explicitly to what Turtle did. The original word "why" is less explicit and could refer to other reasons Turtle won rather than to the means by which it was actually accomplished, (e.g. Rabbit was careless).

(22) You ^ (can / are) run fast, Rabbit. I can't run fast.

In (22) the two sentences are meant to refer to the inherent running abilities of Rabbit and Turtle. The second sentence makes a generic statement about the rabbit's running ability by using the modal "can". However, in the first sentence, because the modal is absent, the statement can be interpreted as either a description of the speed of the rabbit in the race or as generic statement about the rabbit's running ability. The editing miscues make the sentence into a generic one by adding the modal or by using the copula. This makes the first sentence consistent with the second sentence and with the usual way of talking about abilities.

(23) The truck came down ^ (it landed / into) in the hole.

In (23) the editing miscue "it landed" makes the trucks destination more explicit. The substitution of "into" for "in" conveys the idea a little more

explicitly that the truck fell into the hole.

Correcting awkward locutions. Primerese contains many awkward structures. Editing miscues often change these awkward locutions into more natural written prose. In the following examples the form or language of the text as read is changed by the editing miscues to make it conform to more appropriate written prose.

Look at
(24) Look what I can do.

there
(25) The hole was still there. The bump was still up.

In (24) the more natural fixed phrase "Look at" is substituted for the awkward wording "Look what".

In (25) the pragmatically awkward description of the existence of a bump in the road as being "up" produces a miscue that makes the sentence more pragmatically appropriate. This miscue could also be, in part at least, a repeated structure miscue in which the reader expects the word "there" to be repeated.

Another type of editing miscue which changes a primerese feature to a more natural written language form involves dialogue carriers, as can be seen in (26) and (27). The miscue changes the form from the oral and primerese "NP said" to the form more characteristic of written language.[1] We will show below that the primerese form can create ambiguity about the identity of the speaker or addressee.

said the woman
(26) "What! the moon." the woman said.

said the man.
(27) "See! I like my ship," the man said.

Editing miscues at the discourse level. Miscues also show readers' sensitivity to discourse level features.

(28) You run fast Rabbit. I can't run fast.

stopped didn't
But you stop, and I don't.

In (28) the miscues substitute the past tense of the verbs in both sentences, thereby changing the sentence from a generic statement to a

---

[1] NP in NP said stands for noun phrase and includes but is not limited to proper nouns (Lion said); pronouns (he said); articles (a, the) plus nouns (a man said, the man said); and determiners plus adjectives plus nouns (The old man said).

narrative past tense description of what happened in the particular race described in the story.

In (29) the indefinite article is inappropriate because the speaker is referring to a specific hole which has been mentioned in a preceding sentence. The reader's editing miscues change the reference to a definite one which is more cohesive.

this
the
(29) "Just look! I am down in a hole."

(30) 1. Lion said, "Hello, Mouse
2. I want something to eat.
3. I will eat you.
the
4. Don't eat me, said ^ Mouse

In (30) #4, instead of interpreting "Mouse" as a proper noun, which does not take an article, the reader interprets is as a common noun, which can take an article. Thus its second mention requires the definite article to make it cohesive in the discourse, and the reader makes this insertion miscue.

the
(31) One day a fast truck came down a road.

In (30) the indefinite reference "a", which is usually used for new information, is changed to definite reference. The definite article that the reader substitutes in (31) is more appropriate because its referent is identifiable as part of a standard situation which can be assumed to be part of the reader's general schema. This use of definite reference is exophoric and homophoric (Halliday & Hasan, 1976) in that it "is identifiable on extralinguistic grounds no matter what the situation." This type of reference is typical in story openings.

the
(32) He can't have his ship.

In (32) the miscue changes the possessive to the definite article. In this story, "Bad Wish," the fisherman uses one of the his wishes to get a ship. The sentence refers to the ship he asked for, rather than a ship he possessed before the story. Thus the miscue corrects the sentence and makes it more discourse appropriate.

These miscues involving articles suggest that children are not simply substituting one function word for another. Rather they provide more evidence of children's sensitivity to linguistic structure at the discourse level.

Editing miscues influenced by illustration. Illustrations are used in basal texts to carry information that is not explicitly stated in the texts. Thus primerese stories are more dependent on illustrations for their interpretation, i.e. they are picture dependent (Simons & Elster, 1985). Some editing miscues, by putting more information in the text, make the sentences more picture independent, in that they are more interpretable on the basis of the text alone. Other miscues edit the text to make it more consistent with the illustrations.

The phenomenon of adding information to the text that is only provided in an illustration can be seen in (33). The illustration shows a truck up in the air over a large mound in the road. The text and the illustration suggest that the truck flew up into the air after hitting the bump. The text states only that the truck went "up". The editing miscue puts the position of the truck into the text.

in the air.
(33) The bump made the truck go up.^

In the primerese version of the "Lion and Mouse" story the fact that the lion is caught in a net is not stated in the text but is shown in the illustration. The editing miscues in (34) and (35) put more of this information into the text by stating that Lion wants to get "out of" a net which is pictured in the illustration, rather than "away" from a place.

out of
(34) I want to get away from here.

out of
(35) I'll help you get away from here.

This same phenomenon can be seen in (36). The text "Bad Wish" is accompanied by an illustration in which the fish is very large -- almost the size of the boat in the story. The reader inserts the word "big," making the sentence more explicit with respect to the illustration. An alternative explanation of this miscue is that it is influenced by the reader's expectation that the same phrase will repeat itself.

(36) A big fish.

big
What a ^ fish!

In "Look Out" a man is pictured standing on his truck, presumably after getting out of it when it landed in the hole, as stated in (37). The editing miscue in (37) #2 suggests that the reader assumes the man got down from his truck before he uttered (37) #3, rather than uttering it while standing on the truck, as might be suggested by the illustration. This is a plausible interpretation made possible by seeing the illustration as representing one point in time in an ongoing process. The miscue shows that the reader interprets the illustration differently than the author and illustrator of the

text intended. The reader modifies the text through the editing miscue and tries to make it conform to his or her own interpretation of the illustration.

(37) 1. The truck came down in the hole.

off
2. The man got out of his truck.

3. "Look! Look at this he said. ..

Another example of an editing miscue influenced by an illustration can be seen in (38). The illustration shows a lion grasping a mouse in it's paws. In this situation the mouse might conceivably be shouting for help. The reader, upon encountering a word that begins with an "h," makes an illustration-influenced semantic miscue and substitutes the word "help" for "Hello," thus making at least one part of the text more consistent with the reader's interpretation of the illustration. The wrong character is calling for help, so the miscue is not totally consistent with the illustration. The very close graphic similarity between the two words may also have influenced this miscue.

Help
(38) Lion said, "Hello, Mouse ...

Repunctuation Miscues

Another type of miscue mainly produced in the reading of primerese is the repunctuation miscue. Repunctuation miscues include all cases in which the observed response deviates from the expected response with respect to punctuation. Sentences are in effect repunctuated. The intonation patterns the reader produces deviate from intonation patterns predicted by the punctuation. When a period signals the end of a sentence, the expected intonation is stress on the last word, falling intonation, and a pause or juncture. Repunctuation of periods often involves failure to produce the expected intonation. This includes level or rising intonation instead of falling intonation, failure to stress the last word, and failure to pause. An alternative repunctuation miscue at periods in the text involves the production of stress on the last word and pausing, as would be expected, but the maintenance of level or rising intonation instead of falling. Periods are also added to texts by the production of falling intonation, stress, and juncture where there is no period in the text. Failure to produce falling intonation where a comma in the text calls for it was also included as a repunctuation miscue. And, finally, substitution of question intonation or period intonation for exclamation points, and substituting of period intonation for question marks were also counted as repunctuation miscues. Repunctuation miscues demonstrate readers expectations about sentence structure, speech acts, dialogue carriers, and the appropriate oral rendition of written forms.

Truncated sentences. Short, simple sentences 4 to 7 words in length permeate primerese prose. They are written to meet the constraints of readability formulas, which hold that shorter sentences are easier to read

than longer ones. A common repunctuation miscue is the failure to produce stress on the last word and falling intonation as is usually found at the ends of sentences that end with a period. Juncture (pause) is also usually found at the end of a sentence. However this feature is often retained at the end of truncated sentences, but the pause is shortened as in comma intonation. A common feature of these truncated sentences is that they end with verbs (run, stop) or ellipted forms which set up an expectation that other sentences elements are to follow. In (39) through (40) the sentences end with verbs that usually take complements such as prepositional phrases.

(39) You and I will RUN. We will run to the park."

(40) Rabbit said, "I want to STOP. I will stop here.

In (41) and (42) the sentence ends with an ellipted form.

(41) I can run, but Turtle CAN'T. I can get to the park fast.

(42) The truck got OUT. But the hole was still there.

Some subjects' failures to produce sentence ending intonation on truncated sentences suggest they have expectations that more sentence elements will follow. In the reading of primerese, such miscues on ellipted sentences may also reflect expectations regarding the use of repeated structures(See below).

The expectation of more elements to follow can also be seen in (43), which represents a case where the semantics of the verb (use) require a result element (to fill in the hole) as well as an instrument (bump). Subjects showed their expectation of a result element to follow by their failure to produce sentence final intonation and by the insertion of the preposition "to."

to

(43) I will use the bump. I will put it in the hole.

In the rewritten version of the same text a single sentence with a complement clause replaced the two sentences (44). This modification eliminated the repunctuation miscues from the subject's readings.

(44) I will use the bump to fill in the hole.

In some cases readers substitute a lexical item that completes a

truncated sentence instead of omitting sentence final intonation as can be seen in (45).

(45) But the hole was still there.

there.
The bump was still up.

There are other places where sentence-final intonation is omitted between short sentences, but these sentences do not end with obvious linguistic elements which set up expectations that more sentence elements are to follow, as can be seen in (46) through (50). It may be that the sentences are so short and pile up upon one another so that the reader does not feel it necessary to provide the full falling intonation because of the necessity to move on to the next sentence. These miscues may represent an expectations based on the knowledge that written discourse generally involves longer, more complex sentences.

(46) This road has a bump. This road has a hole, too.

(47) He can't have his ship. He can't have his help.

(48) His wish was not good. His wish was too big.

(49) I am down in a hole. I will get my truck out. But you must fix the road. You must make it good.

(50) Rabbit can't see me. I'll get to the park.

Exclamation points and question marks. Other frequently occurring repunctuation miscues are found on exclamation points and question marks. The ubiquitous use of exclamation points in primerese appears to be an attempt to compensate for the omission of verbs and adverbs in dialogue carriers that tell how the quote is to be intoned e.g. "he shouted", "he snarled", "he said excitedly". The "NP said" formula found in primerese does not provide this information. Exclamation points are also apparently used to make the dialogue more animated and to help children "read it like you would say it." Exclamation points are used so profusely in primerese that it is sometimes not obvious how the sentence should be read. Questions marks are more familiar and easier to interpret, and consequently they produced fewer miscues. The dialogues represented in basal readers are unnatural even as spoken discourse, and the repunctuation miscues on exclamation points and question marks appear to be attempts to provide a more natural intonation. The readers repunctuate in order to make the sentences conform to expectations based on knowledge about natural sentence intonation.

The most common miscue on exclamation points is to omit them, especially when they are placed after the first word of a sentence to give it emphasis, as can be seen in (51) and (52). The first two words of these sentences are read as if there was no punctuation and thus no juncture at all is produced.

(51) "Good (!) Now we don't have a bump," they said.

(52) "No (!) fish, I need a big ship," said the man.

Other miscues include the substitution of rising, question intonation for exclamation points as can be seen in (53) and (54). This miscue may be due the fact that children are more familiar with question marks have a better idea of how to intone it. It is easier to give the words questioning intonation that it is to make them exclamations. In (54) this can be seen in the omission of the exclamation point by some subjects and in the substitution of the question by others making the sentence interpretation "Do you see the ship?" rather than "Look at the ship".

(53) "What? The moon?" the woman said.

(54) "You may have a big ship. See?"

Another miscue involved the addition of questioning intonation when it was missing.

(55) See the park?

(56) See what I can do?

In (55) a yes- no question has essentially the wrong punctuation in that the period suggests the kind of intonation associated with a statement. Readers repunctuate the sentence by producing rising intonation which is more appropriate for a yes-no question. (56) is a simple directive "Look at what I can do," but the verb suggests that it is the question form of directive, and some children treat it as one by adding questioning intonation to the sentence.

Primerese Expectations

We turn, now to a consideration of miscues that appear to result from expectations based on the reader's knowledge about linguistic features which are pecular to primerese, as opposed to other, more natural forms of discourse.

Repeated Structure Miscues

As has been discussed, primerese involves the frequent repetition of words in the belief that repetition makes it easier to learn the repeated words. This repetition of words frequently entails the repetition of a whole phrase or most of a sentence. However, because the motivation for the repetition is simply to expose the reader to the words more than once, there is no higher level rhetorical, discourse or story motivation for the repetition. As a consequence it is very difficult for the reader to anticipate when a repetition will occur. In effect, the repetitions in primerese are unpredictable, unlike those in so-called predictable stories found in children's literature, which are both motivated and predictable.

The frequent use of repetition in primerese appears to set up an expectation in readers of primerese that words or structures repeated once will continue to be repeated. The expectation sometimes produces miscues in which words from preceding sentences are substituted in subsequent sentences. In (58) the examples of each individual substitution are from different subjects.

(57) 1. Rabbit said, "I can run.
2. I can run fast.
I can [fa]
3. You can't run,^ Turtle.
4. You can't run fast."
5. Turtle said, "Look rabbit.
6. See the park.
can't
can
7. You and I will run.
I
8. We will run to the park.

In (57) the structure "I can run" is repeated, setting up an expectation that it will be repeated again. This expectation is demonstrated in (57) #3, where one subject substituted "I" for "you". Another subject substituted the modal "can" for "can't". Several subjects produced the partial response [fa]. In (57) #8, the first person singular pronoun "I "is substituted for the first person plural "we". This substitution could be the result of the expectation that the last part of the preceding sentence will be repeated.

The structure pronoun plus modal "can" or "can't" occurs in four out of the six sentences preceding (57) #7, "You and I will run." Thus when a sentence begins with a pronoun there is strong expectation that it will be followed by the modal "can" or "can't," and these modals are substituted in (57) #7. The fact that the source of the substitution does not come from the immediately preceding sentence but appears to come from earlier structures suggests that expectations can be carried over longer stretches of discourse than adjacent sentences.

(58) 1. Mouse said, "No, Lion.
2. I'm not too little to help you.
3. I can help you get away from here."
4. "Do something fast," said Lion.
I'm
I
5. "A man will come and get me. ...

In addition, more general expectations can be built up by the overall pattern in primerese sentences. In the Lion and Mouse story most of the quoted-speech sentences up to (58) #5 begin with the pronouns "I", "I'm" or "you" as in (58) #2 and (58) #3. No sentence begins with determiner plus noun. This pattern of beginning with a pronoun, which is typical of primerese, appears to have built up the expectation that quoted dialogue sentences begin with pronouns that refer to the speaker or the addressee. This expectation is reinforced by the fact that two out of the three preceding sentences begin with a pronoun, and by the fact that the determiner in (58) #5 has only one letter and thus is graphically consistent with the pronoun "I". The result is the substitution of "I" or "I'm" for the determiner "a". This type of substitution occurred for more than 50% of the subjects, suggesting that it reflects a very strong general expectation.

Further evidence for general expectations comes from miscues like the one in (59). The insertion of "little" before "girl" is produced by the expectation that primer stories are peopled by "little N's," which are like fixed phrases. There is also a previous reference to the girl in the story as "a little girl."

little
(59) "Yes I will," said the ^ girl.

Evidence for another kind of generalized expectation comes from miscues related to the repetition of single words, which is another feature of primerese, especially as part of exclamations in dialogue. "Help! Help!," "No! No!," and "Stop! Stop!" all appear in the stories used in this study. The words that are repeated in primerese are probably a limited set (yes, no, help, look, stop). Readers of primerese may develop expectations that these words will be repeated. This expectation can be seen in (60) where the reader treats the second "look" as a repetition of the first "look" and makes a repunctuation miscue by inserting a period with sentence final intonation, so that the phrase "Look at" is broken up and the primerese repetition pattern is maintained. In the earlier part of the story there has only been a single word repetition "Stop! Stop!," this suggests that the miscue in (60) is not produced by an expectation that a specific word will be repeated, but by an expectation derived from a general schema that children have about primerese language.

(60) "Look! Look ^ at this", He said.

(61) 1. He can't have his ship.
2. He can't have his help.

He
3. His wish was not good.

In (61) the repetition of the 3rd person pronoun in (61) #1 and (61) #2 sets up the expectation that the repetition will continue resulting in the substitution of "He" for "His" in (61) #3. The graphic similarity of the printed text and the substituted word increases the probability of substitutions of this type. It might be argued that such miscues are randomly choosen from a set of words that begin with the same letter and are similar in length to the word in the page, and that the role of expectations is minimal. We would argue, however, that while the graphic features of the text do exert an influence, they are subordinate to the structural expectations of the reader. First, in the substitution of function words, while the first letter is often the same, as in (58) above, it can also be different, as in the substitution of "a" for "the", "I" for "he" or "you" etc. It is the category of the substitution that determines the miscue: a pronoun is substituted for a pronoun and a determiner for a determiner etc. Second, the role of discourse expectations that definite reference is required are very clear in (29) where there is no graphic similarity between ""a" and "the" or "this. Finally there are insertion miscues in where a function word added as in (30) #4. In these cases there is obviously no graphic feature at all to constrain the miscue.

The examples of miscues discussed above, in which readers appear to expect words and structures to repeat themselves, suggest that when beginning readers encounter primerese texts, they invoke a schema of primerese language which has developed from repeated exposure to primerese. But repetition expectations are unreliable because of the unpredictability of the repetitions in primerese text, and thus miscues are produced. These miscues almost always result in contextually unacceptable sentences that interfere with the general fluency of the reading.

Miscues Involving Dialogue Carriers

A second structure which produced a number of repunctuation miscues was that of dialogue carriers. In written prose dialogue carriers that enclose quoted speech can take the form of "said NP" or "NP said". While both forms can used the "said NP" for is used more frequently and is considered to be the written form. In oral language "NP said" is more common. Children's literature conforms to more natural written language conventions by using "said NP" most of the time. In primerese, on the other hand, there is a strong tendency for the dialogue carriers surrounding quoted direct speech to take the form of "NP said". In a given story these dialogue carriers may all appear either at the beginning or at the end of quotes. This practice appears to be motivated by the desire to repeat the same words, word order and sentence structure. Placing the "NP said" at the beginning of sentences may also reflect an attempt to ease the load on short-term memory by identifying the speaker at the beginning of the quote. It also appears to be an attempt to mimic oral language.

The use of "NP said" at the end of quoted dialogue sentence or at the beginning of a second quoted dialogue sentence can cause confusion as to which sentence the dialogue carrier should be attached to if punctuation and capitalization is ignored. There is no such ambiguity with the "said NP" form

because it provides a redundant word order cue, in addition to punctuation, about which sentence to attach the dialogue carrier. Beginning readers often make repunctuation miscues on "NP said" dialogue carriers. These repunctuation miscues appear to be a consequence of the over use of "NP said" in primerese. They have the effect of attaching the dialogue carrier to a sentence adjacent to the one intended, so that the quote is attributed to the wrong speaker.

"NP said" can follow a quote or begin a quote. Thus a reader who is processing a few words at a time can easily attach the dialogue carrier to the wrong quote if punctuation is ignored, which often happens. The primerese form "NP said" sets up the expectation that this dialogue carrier can be attached either to the end of one sentence or to the beginning of the next sentence. This expectation often overrides the punctuation cues and results in a repunctuation miscue suggesting that beginning readers are more dependent upon redundant cues than are more mature readers.

(62) Lion said, "Help! Help! I want to get away from here,

but I can't. Mouse said, "Here I am Lion. I'll help you ....

In (62) "NP said" is at the beginnings of the sentences. The period is ignored at the end of the first sentence and is inserted after "Mouse" or after "said". Inserting the period after "Mouse" makes it a part of the quote in the first sentence and indicates the addressee explicitly. Inserting it after "said" attaches the dialogue carrier to the first sentence and suggests that Mouse is the speaker of the first sentence.

(63) "We have what we need," the woman said. No! Fish, I need a big ship," said the man.

In (63) the NP said is at the end of the sentence. In this case the period is inserted at the end of the first quote, the period at the end of the first sentence is omitted and a period is inserted after fish thus placing the dialogue carrier at the beginning of the second sentence. This set of repunctuation miscues attributes the wrong quote to the woman speaker.

Another case of punctuation miscues on dialogue carriers occurs when the dialogue carrier is in the middle of two quoted sentences in which the speaker is the same for both quotes.

(64) "We have help on the ship," the woman said. "Let the fish go."

In (64) a period is inserted at the end of the first quote and the period at the end of the dialogue carrier is replaced by a comma. Thus the dialogue carrier is attached to the second sentence. However, it appears at first glance that attaching the dialogue carrier to the wrong sentence would not be problematical because the same speaker produced both sentences. However, attaching the dialogue carrier to the beginning of the second sentence may be taken to indicate a change of speaker, because this is one traditional way of

indicating a shift in speaker. Thus the reader may attribute the first quote to a different speaker than the second.

Repunctuation miscues with the "NP said" form of dialogue carrier appeared on every instance of this form with the exception of those occuring in two types of context. One was a page-initial "NP said," where the temporal and physical separation of consecutive quotes blocked punctuation miscues. End-of-line breaks did not block dialogue carrier repunctuation miscues. The second context where repunctuation was blocked occurred in cases where the dialogue carrier was preceded by an initial conjunction and where the previous sentence ended in "said NP as in (65).

(65) "Help! Help!" said Lion. Then Mouse said, "Here I am, Lion. I'll help you get out."

When the written language form "said NP" occurred, always at the end or in the middle of quotes, dialogue carrier repunctuation miscues never occurred. In both (66) and (67), which are fairly typical there were no dialogue carrier repunctuation miscues.

(66) "Yes I'm little," said Mouse. But I'm not too little to help you.

(67) "Run fast, Mouse," said Lion. "I'll surprise you" said Mouse.

The sentence structure does not allow the attachment of "said NP" to the following sentence because it cannot be used in sentence initial position. Because "NP said"can appear at the beginning or end of sentences it is possible to attach it to the wrong sentence. Readers appear to need redundant cues as well as punctuation cues to avoid repunctuation miscues. The written language dialogue carrier form "said NP" provides the word order cue in addition to the punctuation cue, and consequently eliminates repunctuation miscues on dialogue carriers.

Educational Recommendations

The results of this study, along with the accumulated literature, lead to several recommendations concerning the texts used to teach beginning reading.

Recommendation # 1: Relax or eliminate vocabulary and sentence controls on beginning reading texts.

The results of this study add to the growing body of evidence that the vocabulary and sentence length constraints currently placed on the reading texts used to teach beginning readers are unnecessary and probably harmful. It has been shown that texts written in more natural language in which these controls are relaxed are not any harder and often are easier to read than the controlled texts. In addition the language of these texts produces unproductive miscues. We therefore would recommend that, within the limits of common sense and guided by the research on children's expectations as shown by reading miscues (see recommendation # 4), vocabulary and sentence length controls be dropped as criteria in the writing and selection of texts used to

teach beginning reading.

Recommendation # 2: Make beginning reading texts less picture dependent.

A majority of the comprehension test items that were the strongest discriminators between the original and rewritten versions were picture-dependent items. Subjects comprehended aspects of the text that were related to the pictures better in the rewritten version, which was picture independent, than in the original version which was picture dependent. In addition, the picture related miscues discussed earlier suggest that the relationship of the pictures to the text can be problematic at times. This leads us to the recommendation that beginning reading texts be made less picture dependent.

Recommendation # 3: Use children's literature in general, and predictable texts in particular.

The vocabulary and sentence length controls imposed on beginning reading texts result in unnatural, primerese language which, as has been shown in this study, produces unproductive reading miscues and interferes with comprehension. We believe that the major criteria for the writing and selection of texts should be the naturalness of the language and the potential interest of the content. Children's literature appears to meet these criteria. Its content, which is often folkloric, is of interest to children. The language provides a better model of natural written language than does primerese. Motivated repetition, particularly in predictable texts, eases the word recognition burden by increasing top-down processing. This repetition also provides the added exposure that makes word learning easier. For all these reasons we recommend that children's literature in general, and predictable texts in particular, be used to teach beginning reading.
Recommendation #4: Use children's miscues as an aid in selecting and writing beginning reading texts.

Not all children's literature is the same. The content varies in interest and the language varies in, among other things, the degree to which it presents a model of natural written prose. The same is true of stories written expressly for children's readers. As Goodman has shown, and as this study has confirmed, miscues provide a window on the reading process. Reading miscues can be used to identify language structures that present problems for beginning readers, e.g. "NP said" dialogue carriers and unmotivated repetitions. These structures can be avoided in the writing of texts, and their frequency of occurrence used as one of the criteria for selecting from among the large amount of children's literature available. The use of miscues as a way of identifying problematic structures is particularly important because it defines natural written language from the child's point of view. That is, it provides information about children's expectations for written prose, and these expectations may be different from those of mature readers. Thus adult models of natural written prose may not be totally appropriate for children at different stages of reading acquisition. The study of children's miscues affords us a method for dealing with this discrepancy. Thus, we recommend the use of miscues as an aid in selecting and writing beginning reading texts.

Conclusions

On the whole, the results of our study indicate that our rewritten versions of basal reader selections were not any harder for first graders to read than the original, "primerese" versions, and in some respects they were easier. Key findings in support of this conclusion are--on the one hand--the virtual absence of differences between versions in reading rate and in overall miscue rate, and--on the other hand--the presence of differences favoring the rewritten versions in the proportion of miscues that were semantically miscues appropriate and in the rate of punctuation. These findings suggest that the rewritten texts are more in line with expectations that children are likely to derive from their knowledge of naturally occurring language patterns.

There is also evidence in our results that the rewritten versions encouraged readers to make greater use of top-down processing and to coordinate different parts of a text, at least among those children who were more able to recognize words from the stories when they appeared in isolation on the word recognition test. Key findings here are the fact that the revised version produced higher comprehension scores within the high word recognition group, along with significant interactions indicating that the rewritten version reduced the proportion of graphically based miscues for high recognition subjects while, if anything, increasing it for low recognition subjects. Related to this is the finding that responses to comprehension questions were especially likely to improve in cases where the essential information was carried by pictures in the original version but was made explicit in the text of the rewritten version. Thus the rewritten versions may also support efforts to base comprehension on the text itself, and not on extra-text sources such as illustrations.

It should be noted that the present study provides a very conservative test of the hypothesis that primerese is harder to read than more natural text. That is, the study was carried out under conditions that almost certainly were less than optimal for showing significant differences in favor of more natural text. For one thing, we did not compare primerese texts with good children's literature. That was because we wanted to hold content constant by comparing "real" primerese stories to the same stories when they were rewritten to make them more natural. Second, our revisions of the original texts were based on intuitions and linguistic analyses reflecting an adult perspective on what constitutes natural language, and they were not guided by data on the ways in which children's expectations might differ from those of adults. In fact our qualitative analysis of miscues has shown that there were points at which our rewriting misfired, as some of the locutions we employed apparently seemed less natural to the children than they did to us. In addition, the quantitative data suggest that, in rewriting one story ("Rabbit and Turtle"), we probably introduced too many new words that were difficult for many of the children to decode.

There were also some problems in the ways we assessed the effects of text differences on decoding and comprehension. The fact that reading miscues occurred so infrequently (in the neighborhood of 5 per 100 words) meant that our quantitative miscue analyses were based on small samples of miscues, making it hard to detect significant differences between the versions in the

relative frequency with which different types of miscues occurred. With respect to comprehension, we know that version differences may have been reduced by ceiling effects on some of the comprehension questions, and it is possible that they were further reduced by the fact that we tested comprehension only after the second reading of each story.

Finally, some of the children who served as our subjects may not have been as well prepared to profit from the rewritten versions as they could have been. For those children who just barely passed the cutoff on our word recognition test, either version of a given story may have still have been too difficult for them to read as continuous discourse, making it virtually impossible for us to show version differences in their reading. Moreover, all of the children had received a steady diet of primerese during several months of reading instruction, but most of them probably had not had as much exposure to more natural written language--certainly not in the context of reading aloud at school.

In spite of all these obstacles, the fact remains that the present study has produced a substantial body of evidence showing that primerese can make reading more difficult for children, and showing how it does so.

REFERENCES

Anderson, R.C., Spiro, R.J. & Montague, W.E. (1977). Schooling and the Acquisition of Knowledge. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Beck, I.L., McKeown, M.G., Omanson, R.C., Pople, M.T. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. Reading Research Quarterly, Spring, 19 (3), 263-277.

Beck, I.L., Omanson, R.C., McKeown, M.G. (1982). An instructional redesign of reading lessons: Effects on comprehension, Reading Research Quarterly, International Reading Association, 17:4.

Beck, I.L., M.G. McKeown and E.S. McCaslin (1981). Does reading make sense? Problems of early readers, The Reading Teacher, April, 780-785.

Brennan, A.D., C.A. Bridge, P.N. Winograd (1986). The effects of structural variation on children's recall of Basal reader stories, Reading Research Quarterly, Winter, 21:1: 91-104.

Bridge, C., Winograd, P. & Haley (1983). Using predictable materials versus pre-primers to teach beginning sight words. The Reading Teacher, 35: 884-891.

Bruce, B., Rubin, A. & Starr, K. (1981). Why readability formulas fail. Reading education report 28, Champaign, Illinois: University of Illinois, Center for the Study of Reading.

Chambers, R. C. (1987). Repunctuation: an exploratory study of text construction in reading. Unpublished dissertation, University of California, Berkeley.

Chall, J. (1958). Readability: An appraisal of research and application. Columbus: Ohio University Press.

Clymer, T. & Parr, B. (1976). "The Lion and the Mouse." In, Helicopters and Gingerbread, (pp. 62-68). Lexington, MA: Ginn & Company.

Clymer, T. & Parr, B. (1976). "Rabit and Turtle." In, A Duck is a Duck, (pp. 62-68). Lexington, MA: Ginn & Company.

Davidson, A. (1984). Readability Formular and Comprehension. In G.Duffy, Roehler and J. Mason (Eds.) Comprehension Instruction, (pp. 128-143). New York: Longman.

Elster, C. & Simons, H.D. (1985). How Important are Illustrations in Children's Readers? The Reading Teacher December, 148-152.

Feldman, M.J. (1985). Evaluating Pre-Primer Basal Readers Using Story Grammar. The American Research Journal, Winter, 22 :4:527-547.

Fry, E. (1968). A readability formula that saves time. Journal of Reading, 11:513-516.

Gates, A.I. (1931). Interest and Ability in Reading. New York: The Macmillan Company.

Gates, A.I. & Russell, D.H. (1938a). Types of materials, vocabulary burden word analysis, and other factors in beginning reading I. Elementary School Journal. September, 27-35.

Gates, A.I. & Russell, D.H. (1938b). Types of materials, vocabulary burden word analysis, and other factos in beginning reading II. Elementary School Journal. October, 119-128.

Gates, A.I. (1961). Vocabulary control in basal reading material. Reading Teacher, 15:81-85.

Gates, A.I. (1962). The word recognition ability and the reading vocabulary of second- and third-grade children. Reading Teacher, 15:443-448.

Goodman, Kenneth S. (1986). For the Reading Commission, NCTE, Basal Readers: A Call for Action, Language Arts, 63:4:358-363.

Goodman, K. (1973). Miscues: window on the reading process. In K. Goodman (ed), Miscue Analysis: applications to reading instruction. Urbana, Illinois: National Council of Teaching of English.

Goodman, K. & Goodman, Y. (1977). Learning about psycholinguistic processes by analyzing oral reading. Harvard Educational Review, 43:317-333.

Gourley, J.W. (1984). Discourse Structure: Expectations of Beginning Readers and Readability of Text. Journal of Reading Behavior, 14:3:169-188.

Gourley, J.W. (1978). This basal is easy to read--or is it?, The Reading Teacher, November, 174-182.

Halliday, M.A.K. & Hassan, R. (1976). Cohesion in English. London: Longman.

Hershman, M. (1984). "Look Out." In, Matteoni, L., Lane, W., Sucher, F., and Yawkey, T. Skylights, (pp. 55-58). Indianapolis, Indiana: The Economy Company.

Klare, G.R. (1963). The Measurement of Readability. Ames, Iowa: Iowa State University Press.

Raskin, L. (1984). "Wish on a Fish," and "A Bad Wish." In, Matteoni, L., Lane, W., Sucher, F., and Yawkey, T. Skylights, (pp. 34-38). Indianapolis, Indiana: The Economy Company.

Russell, D.H. (1938). Types of materials, vocabulary burden word analysis, and other factors in beginning reading II. Elementary School Journal, October.

Samuels, J.J. (1970). Effects of pictures on learnaing to read, comprehension and attitude. Review of Educational Research, 40:397-407.

Schank, R. & Abelson, R. (1977). Scripts, Plans, Goals and Understanding. New Jersey: Lawrence Erlbaum Associates.

Shuy, R.W. and D.L. Lurkin (1978). Linguistic considerations in the simplification/clarification of insurance policy language. Discourse Processes, I:3:305-321.

Shuy, R.W. (1981). Four misconceptions about clarity and simplicity, Language Arts, 358:3:557-561.

Simons. H. & Chambers, R. (1981). Reading error protocol study. Resources in Education.

Simons, H. D. & Leu, D. (1987). The use of contextual and graphic information in word recognition by second, fourth, and sixth grade readers. Journal of Reading Behavior, 19:33-47.

Smith, F. (1982). Understanding Reading. New York: Holt, Rinehart & Winston.

Stanovich, K. (1980). Toward and interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16:32-71.

Weber, R.M. (1970). First graders' use of grammatical context in reading. In H. Levin and J.P. Williams (eds.), Basic Studies in Reading. New York: Basic Books.

Wong Fillmore L., Ammon, P., Ammon, M.S., Delucchi, K., Jensen, J., McLaughlin, B., & Strong, M. (1983). Second year report: Learning English through bilingual instruction. (Contract no. 400-80-0030). Washington, DC: National Institute of Education.

Appendix A
TEXTS AND COMPREHENSION QUESTIONS

RABBIT AND TURTLE

ORIGINAL VERSION

Rabbit said, "I can run. I can run fast. You can't run, Turtle. You can't run fast." Turtle said, "Look, Rabbit. See the park. You and I will run. We will run to the park." Rabbit said, "I want to stop. I will stop here. I can run, but Turtle can't. I can get to the park fast." Turtle said, "I can't run fast. But I will not stop. Rabbit can't see me. I'll get to the park." Rabbit said, "Turtle! You are here! I can run fast, and you can't. But you are here. This is not like you, Turtle." Turtle said, "I do not stop. You run fast, Rabbit. I can't run fast. But you stop, and I don't."

REWRITTEN VERSION

One day Rabbit said to Turtle, "I can run fast, but you can't. I can run faster than you. Turtle said, "Look, Rabbit. Do you see the park? Let's run to the park. We'll see who gets there first." So Rabbit and Turtle began to run. Soon Rabbit stopped to rest. "Turtle can't run fast," he thought. "So I'll take a nap, and I'll still get to the park first. Turtle did not run fast, but he did not stop. He passed the sleeping Rabbit and got to the park. Then Rabbit woke up and ran to the park, but Turtle was there. Rabbit said, "Turtle! How did you get here so fast?" Turtle said, "You can run fast, Rabbit, and I can't. But you stopped and I didn't. That's why I got here first."

COMPREHENSION QUESTIONS

1. What did Rabbit and Turtle decide to do?
2. Why did the animals decide to run?
3. Where did the animals run to?
4. What did Rabbit do after they started to run?
5. Why did Rabbit stop?
6. Why did Rabbit think it was OK to rest?
7. Who got to the park first?
8. Why?
9. Who could run faster?
10. Which of the animals was foolish? Why?
11. Which of the animals was wise? Why?

THE LION AND THE MOUSE

ORIGINAL VERSION

Lion said, "Hello, Mouse. I want something to eat. I will eat you." "Don't eat me," said Mouse. "Let me go, Lion. Let me go. And I'll do something for you." Lion said, "You can't help me. You're too little." "Yes, I'm little," said Mouse. "But I'm not too little to help you. Let me go, and you'll see." Lion said, "I'll let you go. Run fast, Mouse." "I'll surprise you," said Mouse. "You'll see what I can do." Lion said, "Help! Help! I want to get away from here, but I can't." Mouse said, "Here I am, Lion. I'll help you get away from here." Lion said, "You can't help me. You're too little." Mouse said, "No, Lion. I'm not too little to help you. I can help you get away from here." "Do something fast," said Lion. "A man will come and get me. He'll get you, too." Mouse said, "Look here, Lion. See what I can do." Lion said, "You did help me! You are little, Mouse. But you did something big for me." "Yes, I did," said Mouse. "But you did not eat me, Lion. You did something for me, too.

REWRITTEN VERSION

One day a Lion caught a Mouse and said, "Hello, Mouse. I want something to eat, so I will eat you." "Don't eat me," said Mouse." Let me go, Lion. If you let me go, I'll do something for you some day." "You can't help me," said Lion. "You're too little." "Yes, I am little," said Mouse. "But I'm not too little to help you. Let me go, and you'll see." Lion said, "I'll let you go." And he did. "Run fast, Mouse," said Lion. "I'll surprise you," said Mouse. "You'll see what I can do." When Mouse met Lion again, Lion was in a net and he could not get out. Help! Help!" said Lion. Then Mouse said, "Here I am, Lion." I'll help you get out." "You can't help me," said Lion. "You're too little." "No, I'm not too little," said Mouse. "Then do something fast," said Lion. "A man is coming to get me, and he'll get you, too." "Look here, Lion," said Mouse. "Look what I can do." Mouse nibbled a hole in the net and Lion got out. Then Lion said, "You did help me, Mouse! You are little, but you did something big for me." "Yes, I did," said Mouse. "But you did something for me, too. You didn't eat me."

COPMREHENSION QUESTIONS

1. What did Lion do when he first met Mouse?
2. What did Lion want to do?
3. Why didn't Lion eat Mouse? Probe (What did Mouse say so that Lion wouldn't eat him?)
4. Did Lion believe that Mouse could help him?
5. Why?
6. Why did Lion Say, "Run fast, Mouse?"
7. What was happening when Mouse met Lion again?
8. Why did Lion want help?
9. What did Mouse do to help Lion?
10. What would have happened if Lion didn't get out of the net?
11. Why did Mouse help Lion?

LOOK OUT!

ORIGINAL VERSION

The people came to the road. Then a fast truck came. "Stop!" the people said. "Stop! Stop! This road is not good. This road has a bump. This road has a hole, too. Go slow, truck! Go slow!" But it was too late!
The truck was on the bump. The bump made the truck go up. The truck came down in the hole. The man got out of his truck. "Look! Look at this," he said. "Just look! I am down in a hole. I will get my truck out. But you must fix the road. You must make it good. "The truck got out. But the hole was still there. The bump was still up. "What can we do?" they said. "I will fix it," she said. "What! You will fix the road?" they said. "You can make it good?" "Yes! I can do it," she said. "I will use the bump. I will put it in the hole." "Good! Now we don't have a bump," they said. "That is good!" And we don't have a hole. That is good, too. We have a good road."

REWRITTEN VERSION

One day a fast truck came down a road. "Stop! " said the people. "This road is no good. It has a bump and it has a hole, too. Slow down!" But it was too late! The truck hit the bump and up it went. Down came the truck in the hole. The man got out of his truck. "Look at this!" he said. "Just look at my truck down in this hole. I can get it out, but you should fix the road." The man got his truck out but the hole was still there. And so was the bump. "What can we do?" said the people. "I will fix it," said a little girl. "What! You will fix the road?" said the people. "Yes, I will," said the girl. "I will use the bump to fill in the hole." "Good!" said the people. "Now we don't have a bump and we don't have a hole. We have a good road."

COMPREHENSION QUESTIONS

1. What was wrong with the road?
2. What did the people say to the truck driver?
3. What happened to the truck?
4. What did the truck driver say or tell the people to do?
5. Who got the truck out of the hole?
6. Who offered to fix the road?
7. Did the people believe she could do it?
8. Why? Why not?
9. How did the girl fix the road?
10. What did the people say then?
11. How did they feel when she fixed the road?

A BAD WISH

ORIGINAL VERSION

"Look, a fish!" said the man. "A big fish! What a fish! It will be good to eat. Let's eat it." No! No! Don't eat me," said the fish. I will let you have a wish." "We have what we need," the woman said. No! Fish, I need a big ship," said the man. The fish said, "You may have your wish. You may have a big ship. See!" "I like my ship," the man said. We can let the fish go," the woman said. "We have the big ship." "No! Fish, we must have help on this ship," the man said. The fish said, "You may have your wish. Look!" "We have help on the ship," the woman said. "Let the fish go." "We have a ship," the man said. "We have help on it. But I still need the fish. It must stay!" "No! You don't need the fish," the woman said. "We have what we need. Let the fish go." "No! I need the moon, too," the man said. "Fish! I must have the moon." "What! The moon!" the woman said. "Fish, he can't have the moon, can he?" "No, no!" the fish said. "He can't have the moon! It's too big! It's not a good wish. It's a bad wish. He can't have the moon. He can't have his ship. He can't have his help. His wish was not good. His wish was too big. It's not good to stay. I will not stay."

REWRITTEN VERSION

One night a man and a woman were in a small boat and they caught a fish. "Look!" said the man. "What a big fish! It will be good to eat." "No! Don't eat me," said the fish. "If you don't, I will let you have a wish." "We have everything we need," said the woman. "No, we don't!" said the man. "Fish, I need a big ship." "You may have your wish," said the fish. So the man got a big ship, and he was happy. "Now we have a big ship," said the woman. "So we can let the fish go." "No!" said the man. "Fish! We must have helpers on this ship." "You may have your wish," said the fish. "Look!" And there were helpers on the ship. "Now we have helpers," said the woman. "So we can let the fish go." "No, we can't," said the man. "We have a ship and helpers, too. But I still need the fish." "No, you don't!" said the woman. "We have everything we need. Let the fish go." "No! I need the moon, too," said the man. "Fish! I must have the moon." "What! The moon?" the woman said. "Fish, he can't have that, can he?" "No!" said the fish. "He wants too much. It's a bad wish. He can't have the moon. And now he can't have his big ship or his helpers." So the man and the woman were back in the small boat.

COMPREHENSION QUESTIONS

1. Where were the man and the woman at the beginning of the story?
2. What did the man want to do when he first saw the fish?
3. What did the fish say?
4. What were the three wishes the man made?
5. What did the woman want the man to do?
6. Why?
7. Did the fish let the man have the moon?
8. Why? Why not?
9. Why was the man's wish bad?
10. What happened to the man and the woman at the end of the story?
11. What happened to the fish at the end of the story?

APPENDIX B

WORD RECOGNITION TEST WORD LIST

but how want too rest late fill slow help

got must need bump small girl stay left now

moon there fish have don't truck then park let's

still boat down ship soon here than

your one little again we'll use woke

first hello what I'll faster were happy

helpers coming surprise could something thought you're people

began sleeping caught everything nibbled should woman passed